UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SEVERE HAIL DETECTION USING REMOTE SENSING OBSERVATIONS AND AN EVALUATION OF THE ABOVE ANVIL CIRRUS PLUME FOR SEVERE WEATHER DETECTION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

ELISA M. MURILLO
Norman, Oklahoma
2018

SEVERE HAIL DETECTION USING REMOTE SENSING OBSERVATIONS AND AN EVALUATION OF THE ABOVE ANVIL CIRRUS PLUME FOR SEVERE WEATHER DETECTION

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

Dr. Cameron Homeyer, Chair

Dr. Corey Potvin

Dr. Amy McGovern

Dr. Jeff Basara

Acknowledgments

I would first like to thank my advisor and science dad, Dr. Cameron Homeyer. This work was only possible through his consistent motivation, encouragement, attention to detail, and countless hours of coding help. Dr. Homeyer continually pushes me to think in new ways, teaches me new skills (both for research and life), and makes me a better student, researcher, and overall person through his guidance. Even though I've had several familial obstacles arise during my time here, he always supported and encouraged my decisions to prioritize my family needs, yet he never wavered in his confidence in me to do important and valuable research. He is honestly a role model for many other advisors! While I would prefer less dad jokes, I am willing to pay that small price to have Dr. Homeyer as my advisor!

I also must thank my collaborator, Kristopher Bedka. He has been on top of everything and very willing to work with my schedule, resulting in a publication that we are all very proud of. He also, similar to Dr. Homeyer, has pushed me to be a better scientist, and I thank them both for my growth. I also thank Dr. John Allen for his recommendations regarding hail reports; the results carry more weight now given those suggested changes.

I must also thank Dr. Daphne LaDue for accepting me as an REU participant. During my time as a participant and an assistant director, I met Dr. Homeyer (current advisor), Dr. Corey Potvin (REU advisor and current committee member), and Dr. Amy McGovern (current committee member). Having already knew the majority of my committee when starting my graduate studies, I felt more confident with them all my by side. With that being said, I also must thank my committee members for their advice and support throughout my entire time here at OU. I would also like to thank Dr. Jeff Basara (current graduate liaison and member of my committee) for always having my best interest in mind (and every other student). Additionally, I would like to thank Christie, Shelby, and the rest of the SoM staff for their immense efforts to make sure that all of us students are taken care of.

My research group, tripp-C, current and past, has also been an important part of my grad school journey. In particular, Thea Sandmæl has been like a big sister to me; given that our research is closely related, she took me under her wing to help me throughout the entire graduate school process. I also want to acknowledge several close friends of mine, including Amanda Murphy, Briana Lynch, and Joel McAuliffe. They have all sat through numerous practice presentations, providing awesome feedback! Their support and encouragement have been invaluable, and they provide an ideal example of what friendship should be like.

A very important person deserves their own thank you! My life partner, Brian Robert Greene has been the backbone of my support system, however, he serves many other roles, too: best friend, presentation critic, ally for minorities, paper reviser, greatest Elisa advocate, adventure partner, brainstorm catalyst, and so many more. Thank you for every day we've had together and for every day we will have. Thank you for all of the love and support you give me. I am forever grateful to be an OU graduate student, as that led me to meet the love of my life!

Finalmente, necesito agradecer a mi familia. Los últimos años han estado muy difíciles para todos, pero aun así nunca ha habido una falta de amor, apoyo, y animo de nadie. Mis padres (Mom, Dad, Tracey) y mis hermanos (Gus, Kelsie) han estado allí para mí, todos los días. Sinceramente, sin ellos, no yo estaría donde estoy ahora. También quiero agradecer a Mamia, Papi, Camila, y Miguel por su constante amor y apoyo. Cada uno de ustedes aquí son un modelo a seguir para mí.

Table of Contents

List of Tables

List of Figures

Abstract

This thesis undergoes a comprehensive comparative analysis to address two main objectives: a) Thoroughly assess available radar, satellite, and lightning based products' ability to identify hail events and size, and b) quantify the relationship between AACPs and severe weather. Radar observations have been used in an abundance of studies aiming to identify severe weather hazards. However, due to historical reporting limitations, hail events have not been studied as extensively as other hazards. In the contiguous US, real-time analysis and detection of severe storms is largely conducted using ground-based radar observations, especially those from the operational Next Generation Weather Radar network (NEXRAD), which provides three-dimensional information on the physics and dynamics of storms at 5-min intervals. Recent NEXRAD upgrades to higher resolution and dual-polarization capabilities (completed in 2013) have provided improved capability to discriminate between hydrometeor types in real time. Additionally, new generation geostationary satellite observations (e.g. GOES-16) have greater spatiotemporal resolution than their predecessors, leading to significant changes in the utility of satellite observations.

Improved satellite observations also enable better identification of above anvil cirrus plumes (AACPs), which have been identified in satellite imagery for decades. AACPs are generated from intense tropopause penetrating updrafts and gravity wave breaking. AACPs are often evident during severe weather outbreaks and, due to their prominence, have been studied for several decades.

Thorough assessment of established products' ability to identify hail occurrence and size has been uncommon. Additionally, current published research indicates that the AACP signature is one of the strongest indicators of a severe storm depicted by visible and IR satellite imagery. Despite the extensive research devoted to the AACP, there is lingering uncertainty regarding how these processes are depicted by datasets routinely used in operations and how AACP recognition can augment severe storm identification.

For hail detection assessment, we compare the most commonly used objective hail identification methods with less common approaches, focusing on quantitative metrics from both satellite and radar observations over the CONUS. We find that H_{DR} below the melting level (BML), maximum estimated hail size (MESH), and vertically integrated liquid (VIL) density provide the best hail event indication, while a combination of either MESH or VIL density combined with H_{DR} BML provide the best hail-producing storm discrimination. The MESH power relation was also refit to the study's report dataset, and outperformed the Witt et al. (1998a) MESH for the maximum estimated hail size.

AACP and non-AACP storms are linked with storms, National Weather Service warnings, and severe weather. AACP storms were shown to be much more likely to be severe than non-AACP storms, and the majority of significant severe weather reports are produced by AACP storms. Additionally, the ability of National Weather Service (NWS) warnings with and without AACP information to detect storms that produced 2+ inch hail was assessed. It was found that the addition of AACP identification markedly increased the discrimination ability of these storms.

Chapter 1

Introduction

1.1 Motivation

Between 1980-2018, severe weather has caused the most billion-dollar disasters out of numerous natural hazards (e.g. drought, flooding, freezes, winter storms, tropical cyclones, and wildfires) (Smith and Katz 2013; Smith and Matthews 2015). Between 2013-2017 alone, severe weather damage accounted for over $16 billion of property and crop damage, with the greatest portion ($\sim$56%) produced by hail (NOAA Severe Weather Database). As a result of its economic impact, severe weather has been studied across the globe since the mid-1900s (e.g. Fujita et al. 1970; Mather et al. 1976; Waldvogel et al. 1979). Many studies have utilized radar-based methods, using both single-polarization (SP) radar observations and dual-polarization (DP) radar observations, to study severe weather events (Cook 1958; Forbes 1981; Klimowski et al. 2003). Radar, satellite, and lightning observations have been used in an abundance of studies aiming to identify hail occurrence and size. However, to date, a thorough assessment of many established products' ability to identify hail occurrence and size has been uncommon, or for some products, do not exist.

One particular cloud feature seen in satellite imagery, known as an above anvil cirrus plume (AACP), is formed when intense tropopause-penetrating updrafts located in strong storm-relative wind shear environments inject ice into the lower stratosphere. These AACPs are often evident during severe weather events and, due to their importance, have been studied for 35+ years (e.g. McCann 1983). Despite this research, there is uncertainty regarding exactly how severe are storms with AACPs, and how AACP identification can assist with severe weather warning.

This study seeks to address two main objectives: a) analyze SP radar, DP radar, satellite, and lightning products to determine the best indicators for identifying hail events and hail-producing storms and b) quantify the relationship between AACPs and convective hazards. For hail detection assessment, we compare the most commonly used objective hail identification methods and other less common approaches, focusing on quantitative measurements over the CONUS to answer the following questions:

1. What are the best measurements to identify hail occurrence?

2. What are the best measurements for distinguishing between storms that produce severe hail and those that do not?

3. Given the historically limited availability of DP radar observations, can SP radar estimates of hail size be improved for future climate studies?

To address our second main objective, AACP and non-AACP storms are collocated with available radar, satellite, and lightning based severe weather identification products and all types of severe weather to answer the next set of questions:

1. What is the severe weather frequency for AACP storms compared to non-AACP storms?

2. Are certain severe weather types more likely to be produced by AACP storms?

3. How can severe weather warnings be augmented by AACP identification?

4. What is the correlation between AACP storms and supercells? What is unique about supercells that do not produce AACPs?

1.2 Background

Numerous objective methods to identify hail occurrence and size have been developed and documented in the literature, some of which are in use operationally and pre-date

the Next-Generation Weather Radar (NEXRAD) network in the United States (e.g. Cook 1958; Petrocchi 1982; Cheng et al. 1985). Radar-based methods generally fall into two groups: those based on SP radar observations and those based on DP radar observations. Notable SP radar methods that have been shown to perform well are the vertically integrated liquid water (VIL) density (Amburn and Wolf 1997) and the maximum expected size of hail (MESH; Witt et al. 1998a), both of which are based on vertical integration of radar reflectivity at horizontal polarization (Z_{H}). These and other similar Z_{H}-based methods have been cross-evaluated and/or verified against hail reports in multiple studies during the past two decades (e.g. Edwards and Thompson 1998; Holleman et al. 2000; Marzban and Witt 2001; Ortega et al. 2005; Ambrosio et al. 2007; Donavon and Jungbluth 2007; López and Sanchez 2009; Liu and Heckman 2010; Cintineo et al. 2012; Skripniková and Řezáčová 2014; Capozzi et al. 2018; Wang et al. 2018; Ortega 2018).

While many studies have found SP radar parameters to be useful for (and skillful at) identifying severe hail-producing storms, methods used for performance evaluation and the sample sizes used for verification have varied considerably. Few studies exist which evaluate these parameters for large sample sizes ($>$1000 severe hail reports), broad regions, or long time periods. Additional limitations exist when considering the challenge of hail size discrimination, for which many SP parameters have shown limited utility. For example, Edwards and Thompson (1998) found that VIL density was not useful for operational hail size discrimination, and Picca and Ryzhkov (2012) found that MESH overestimated hail size. Providing reliable estimates of hail size based on radar observations continues to be a significant challenge.

Though the NEXRAD network was only recently upgraded to DP capabilities, non-operational DP research radars have been used to study severe weather, and specifically hail events for many years. Unique signatures in the DP variables have been shown to offer improvements for identifying particle types, since these observations provide more information on hydrometeor shape, phase, and size. This includes the discrimination between

meteorological and non-meteorological scatterers, which resulted in ways to identify tornadic debris in these observations (tornadic debris signature; e.g. Ryzhkov et al. 2005c). There are additional improvements in hail identification and hail size discrimination, given the new shape/size information (Aydin et al. 1986; Bringi et al. 1986; Kennedy et al. 2001). In particular, one can discriminate between large raindrops and hail stones when using these new observations. Large raindrops fall as oblate spheroids while large hail tends to tumble as it falls and appear spherical in DP radar observations. Thus, collocated DP observations of Z_{H} (proportional to the sixth moment of a particle size distribution) and differential radar reflectivity (Z_{DR}; a size-weighted measure of particle shape) enable improved spatiotemporal discrimination between regions of hail and rain at low levels compared to an approach based on Z_{H} alone. Hail, when large and approximately spherical, is typically characterized by $Z_{\mathrm{H}} \geq 45$ dBZ and $Z_{\mathrm{DR}} \approx 0$ dB, while large raindrops are often characterized by similar Z_{H} and $Z_{\mathrm{DR}} \geq 2$ dB. There have been a number of approaches used to leverage these scattering characteristics for hail identification. One approach has led to the creation of the so-called "hail differential reflectivity" (or $\mathrm{H_{DR}}$; Aydin et al. 1986; Depue et al. 2007), which uses DP observations to identify deviations from the expected relationship between Z_{H} and Z_{DR} for rain. Another approach has been to use "fuzzy" logic to determine the most likely hydrometeor type based on its scattering characteristics using multiple DP radar variables (e.g., Vivekanandan et al. 1999; Straka et al. 2000; Heinselman and Ryzhkov 2006; Elmore 2011; Mahale et al. 2014). This technique is commonly referred to as a hydrometeor classification algorithm (HCA). Park et al. (2009) outline the operational HCA used in the NEXRAD network, which features a rain-hail mixture classification as its largest hydrometeor category. A recent improvement to this operational HCA has been made to distinguish between three hail size categories, which was shown to outperform the current operational SP hail detection algorithm (referred to as the hail size discrimination algorithm; Ryzhkov et al. 2013; Ortega et al. 2016).

While these radar methods are useful when data are available, there are many coverage gaps over the U.S., and many other countries do not have an extensive radar network. Additionally, there are other factors that limit the accuracy of radar observations, including but not limited to increased beam height with distance from radar site, returned power from non-meteorological phenomena, and time delay of full radar volume scan. Therefore, when severe weather identification is solely reliant on radar observations, the potential for detection could be limited.

Characteristics of lightning have been used for updraft identification, though mostly focused on broad severe weather application rather than specific hazards. In particular, lightning flash rates have been used to identify the precise locations of strong convective updrafts, which are often the source of severe weather events. For example, a phenomenon referred to as a "lightning jump", representing a rapid increase in total lightning, has been correlated with severe weather occurrence in many studies (e.g. Williams et al. 1999; Deierling and Petersen 2008; Schultz et al. 2009; Liu and Heckman 2010). Schultz et al. (2017) found lighting jumps useful for identifying strong, mixed-phased updrafts, which subsequently increases the likelihood of falling hail. Farnell et al. (2018) also found that lightning jumps are useful for not only long lead times of approximately 55 min for hail events, but also for locating large ($>$ 2 cm diameter) hail fall.

Over the past several decades, unique patterns within convective cloud tops and cloud evolution depicted by geostationary satellite imagery have been deemed useful by the operational forecasting or research communities for severe storm detection and/or nowcasting. However, satellite observations have received little attention for more specific identification, such as hail occurrence. This limited use is largely due to the fact that geostationary satellite imagery has traditionally been available at coarse spatial and temporal scales relative to radar observations. The Geostationary Operational Environmental Satellite (GOES) systems have been the primary source of US satellite weather observations for several decades, given that two are typically operational at a time. Prior to the launch of

the GOES–R satellite in 2017 (now operationally known as GOES-16), GOES–13, –14, and –15 were operational and featured visible imagery at 1 km horizontal resolution and IR imagery at 4 km horizontal resolution, with image updates every 5 to 15 minutes (Menzel and Purdom 1994). Beginning in 2012 when GOES–13 and –15 filled the operational east and west positions, GOES–14 was operated in Super Rapid Scan Operations for GOES–R (SRSOR) mode that collected imagery at 1-minute intervals in preparation for the increased temporal resolution of the GOES–R series. The new system features the Advanced Baseline Imager which has 0.5 km horizontal resolution for visible, 2 km for IR imagery, and can operate in "Flex Mode" for temporal resolution of 30 seconds to 1 minute (Schmit et al. 2005, 2014). Higher temporal resolution datasets enable better early recognition of satellite features, producing a more complete understanding of their severe weather linkage.

Despite pre-GOES-R limitations, several cloud phenomena have been associated with severe weather occurrence, and more recent studies have utilized the higher resolution datasets to better understand those phenomena (e.g. Cintineo et al. 2013, 2014; Gravelle et al. 2016; Line et al. 2016; Bedka and Khlopenkov 2016). Such phenomena include rapid cloud top cooling during storm initiation (Cintineo et al. 2013), cloud tops extending above the equilibrium level of deep convection referred to as overshooting tops (OTs; Brunner et al. 2007; Setvák et al. 2010; Dworak et al. 2012; Bedka and Khlopenkov 2016), the presence of anomalously small water droplets or ice crystals within developing cumulus and anvils (Lindsey et al. 2006; Rosenfeld et al. 2008), enhanced wind flow divergence and vorticity within anvil cloud top (Apke et al. 2016), anomalous storm motion relative to other nearby storms indicative of a supercell storm (Bunkers et al. 2006), and the so-called "enhanced-V" in infrared (IR) imagery, now understood to represent the occurrence of AACPs (McCann 1983; Brunner et al. 2007; Setvák et al. 2010; Dworak et al. 2012; Púčik et al. 2013; Homeyer 2014; Bedka et al. 2015; Homeyer et al. 2017; Kunz et al. 2018). While these studies documented the qualitative and quantitative benefits of GOES–14 SRSOR observations for severe storm discrimination, there remains a lack of focus on

hail events in particular. Even with the addition of hail-specific studies using GOES–14 SRSOR observations, the specific utility/benefit of GOES–16 (and similar future systems) for enhancing severe (and hail) storm identification for remains mostly unknown due to its novelty.

AACPs are of particular interest in this thesis, as they have been examined for decades showing connections to severe weather occurrence. AACPs are formed when intense tropopause-penetrating updrafts (i.e. OTs) located in storm-relative wind shear environments favoring gravity wave breaking inject ice into the lower stratosphere which sublimate quickly to form cirrus clouds. They can reside several kilometers above the primary anvil (Wang 2003; Setvák et al. 2013; Homeyer et al. 2017) and often cast shadows on the primary anvil near sunset view-able in visible imagery. Even though OTs are required to generate AACPs, only a small subset of OTs, those that penetrate the tropopause >1 km within storm-relative wind shear environments favorable for gravity wave breaking, were found to generate AACPs (Homeyer et al. 2017). The lofting of cloud material during a wave breaking event is consistent with the early physical descriptions of cloud-top evolution during plume formation by Fujita (1982), where it was called "jumping cirrus." The differing ice microphysics and height residence between the AACP and primary anvil combine to generate unique textures that allow AACPs to be readily identified in visible wavelength imagery. In many situations, the AACP adjusts to the ambient temperature of the lower stratosphere, causing the AACP infrared (IR) brightness temperature (BT) to be warmer than the underlying anvil, most notably adjacent to the OT updraft region. The warm AACP BT contrasts sharply with the cold OT that is continuously reinforced by rapidly rising air that cools upon ascent due to adiabatic expansion. However, AACP IR BT patterns are not always warmer than the surrounding cloud (see section 4.1 for a more detailed discussion). Further studies that test different causes of cold AACP signatures are required to fully understand the specific situation of AACP characteristics.

AACPs (or its other aliases) were one of the first satellite features associated with severe weather (e.g. McCann 1983), and recent progress regarding this relationship is underway due to higher resolution satellite imagery. Storms with OTs and no AACP do often generate hazards such as lightning, heavy rainfall, aviation turbulence, and aircraft engine icing conditions (Bedka et al. 2010; Yost et al. 2018). However, severe weather is still less common for OT storms without AACPs. Bedka et al. (2015) manually identified 58 AACP-producing storms using GOES–14 SRSOR observations and showed that 57% of the AACP cells were severe, and that AACPs appeared on average 18 minutes before the storm's first severe weather report. Even though current published research indicates that the AACP signature is one of the strongest indicators of a severe storm depicted by visible and IR satellite imagery, there is lingering uncertainty regarding how these processes are depicted by datasets routinely used in operations and how AACP recognition can augment severe storm identification.

The development and evaluation of methods for objective severe weather detection is not only motivated by an operational need of such products for forecasters, but also for high-confidence verification of numerical models and climate studies. Although the NCEI and SPC databases provide the most comprehensive accounts of historical severe weather events in the U.S., they are plagued by serious limitations (e.g. frequency and magnitude) that can influence the severe weather relationships presented in this thesis (Doswell et al. 2005; Trapp et al. 2005, 2006; Verbout et al. 2006; Brotzge et al. 2011). For hail reports in particular, there are biases in frequency, reported sizes, time of occurrence, and location (Kelly et al. 1985; Bardsley 1990; Witt et al. 1998b; Fraile et al. 1992; Schaefer et al. 2004; Allen and Tippett 2015; Blair et al. 2017; Allen et al. 2017; Witt et al. 2018). Due to safely concerns, hail is often reported after the storm has passed, which commonly biases the time of hail occurrence in the report database. In addition, spatial and temporal distributions of hail reports are influenced by population, road networks, storm chasers and time of

day (Allen and Tippett 2015). For example, a strong storm could produce hail in a low-population region and result in few to no reports. On the other hand, an equivalent storm in a region with a larger population or heavily trafficked road network is likely to result in more frequent and representative hail reports than the previous scenario. Even with sufficient witnesses present, more severe hazards tend to be reported more frequently at the expense of less severe hazards (Kelly et al. 1985; Witt et al. 2018). It is less likely for hail to be reported when a tornado is occurring simultaneously, a result of people 1) taking shelter from the present tornado and therefore unable to observe hail and/or 2) reporting what they perceive as the more threatening hazard. One of the greatest limitations with this database, however, is the error associated with hail sizes. Reporters often use reference objects to estimate hail size, resulting in a clustering of reports at specific sizes. This produces a discretized hail size distribution that may not be representative of the true size distribution.

Even with accurate hail stone measurements, a range of sizes often occurs for a given hail-producing storm, making it challenging to verify particular sizes. Several field campaigns have collected high-resolution hail observations that have advanced our understanding of spatiotemporal variability of hail occurrence beyond that possible with traditional reporting databases (e.g. Strong and Lozowski 1977; Ortega et al. 2009; Blair and Leighton 2012; Heymsfield et al. 2014; Blair et al. 2017), but these data are still limited to relatively small sample sizes. Thus, determining the most reliable method of objective hail occurrence and hail size estimates over large regions using long-term, high-resolution radar or satellite records is needed to advance the understanding and prediction of hail events over a range of weather and climate timescales.

Chapter 2

Data

2.1 Cases

Days to include in this analysis were chosen based on 1) data availability, 2) environmental set up, 3) dominant convective mode, and 4) severe weather frequency/magnitude, resulting in 30 cases that represent a variety of convective scenarios. 16 of the 30 days were chosen due to the availability of 1-minute GOES satellite imagery. Given temporal resolution requirements, only these 16 days were considered when calculating satellite-based cloud-top divergence (Apke et al. 2016). The cases are clustered mainly in the central U.S., but extend throughout most of the east as well. While the majority of the cases are dominated by discrete or semi-discrete convection, seven of the 30 days are dominated by mesoscale convective systems (MCSs). Some cases were chosen due to severe and widespread damage or visual confirmation of severe weather occurrence by a team member. Table 2.1 specifies the locations, storm totals and report totals for all 30 days (storm totals result from objective storm tracking detailed in section 2.8).

While it is possible to distinguish the presence of AACPs in 5- to 30-minute satellite imagery, 1-minute resolution is optimal for identifying AACPs, which can sometimes be subtle and short-lived, and determining the start and end times of plume production. It is also difficult to link AACPs to objective radar storm tracks (described in section 2.8), particularly in MCS cases. Therefore AACP evaluation was limited to the remaining 13 non-MCS, 1-minute imagery cases (Table 2.2). AACP analyses were only completed within the temporal and spatial domains in which 1-minute imagery were available. Since the satellite domain was only a sub-domain of the full analysis domain for seven of the 13 AACP days (italicized in Tables 2.1 and 2.2), only the smaller sub-domain was evaluated during AACP analyses for these cases.

Table 2.1: Details of all 30 days used during analysis. Cases featuring 1-minute satellite imagery are bolded, and those in which this imagery is only available for a sub-domain of the full analysis domain are italicized. Cases featuring GOES-14 SRSO observations, in particular, are denoted with "*".

Date	No. Storms (No. Severe)	No. Hail Storms (No. Reports)	Analysis Domain Coordinates [lon_0, lat_0, lon_1, lat_1] (°W and °N)	AACP availability
***12 June 2013**	555 (132)	36 (380)	[96.0, 38.0, 80.0, 45.0]	–
27 April 2014	223 (34)	32 (110)	[99.0, 34.0, 91.5, 42.0]	–
***10 May 2014**	112 (37)	35 (363)	[99.0, 36.0, 90.0, 43.0]	Yes
***11 May 2014**	330 (58)	48 (398)	[102.0, 36.0, 92.0, 44.5]	Yes
***21 May 2014**	54 (9)	8 (79)	[106.0, 37.5, 101.0, 41.0]	Yes
16 June 2014	406 (36)	31 (150)	[100.0, 41.0, 89.0, 44.0]	–
17 June 2014	155 (20)	17 (98)	[106.0, 41.5, 94.5, 48.0]	–
18 June 2014	79 (6)	2 (9)	[100.0, 43.5, 98.0, 46.5]	–
13 October 2014	707 (610)	11 (55)	[95.5, 29.5, 84.5, 40.5]	–
6 May 2015	202 (44)	32 (115)	[100.0, 32.5, 95.5, 41.5]	–
***19 May 2015**	329 (37)	22 (121)	[103.0, 29.0, 94.0, 37.0]	Yes
***24 May 2015**	123 (11)	11 (98)	[105.0, 36.0, 97.0, 41.0]	Yes
***25 May 2015**	669 (111)	30 (159)	[105.0, 25.0, 89.0, 41.0]	–
***27 May 2015**	387 (48)	43 (392)	[104.0, 29.5, 96.0, 41.5]	Yes
***4 June 2015**	290 (39)	35 (208)	[108.0, 34.0, 93.0, 43.0]	–
23 December 2015	137 (31)	25 (121)	[92.5, 33.5, 84.0, 42.0]	–
15 April 2016	160 (27)	25 (131)	[104.0, 34.5, 99.0, 40.5]	–
9 May 2016	199 (50)	49 (235)	[100.0, 33.0, 94.0, 41.5]	–
24 May 2016	150 (21)	18 (70)	[104.0, 35.5, 97.0, 41.0]	–
25 May 2016	17 (5)	5 (17)	[99.5, 35.5, 95.0, 40.0]	–
28 March 2017	458 (49)	30 (249)	[105.5, 27.5, 92.0, 39.0]	Yes
3 April 2017	499 (74)	12 (58)	[90.5, 29.0, 78.5, 37.0]	–
5 April 2017	496 (119)	78 (776)	[90.0, 30.0, 78.0, 41.0]	Yes
16 May 2017	947 (138)	88 (472)	[105.0, 30.5, 87.5, 48.0]	Yes
18 May 2017	691 (152)	70 (364)	[104.5, 31.0, 89.5, 40.5]	Yes
28 June 2017	604 (82)	37 (133)	[104.0, 36.5, 87.0, 47.5]	Yes
29 June 2017	520 (68)	44 (300)	[105.0, 35.5, 90.0, 45.0]	Yes
19 July 2017	514 (77)	8 (34)	[104.0, 40.5, 87.5, 48.5]	Yes
Total	**10,650 (2,212)**	**956 (5,954)**	–	**13**

Table 2.2: Additional information on the subset of cases from Table 2.1 used for AACP identification, including a detailed storm and report breakdown of each case. Italicized dates indicate those in which the 1-minute imagery is available for only a sub-domain of the full analysis domain outlined in Table 2.1. Significant severe weather is defined as hail reports $\geq$ 2 in, wind reports $\geq$ 65 knots, and/or tornadoes resulting in $\geq$ EF2 damage. Descriptions for all storm, supercell, and AACP identification methods are outlined in later sections.

Date	No. Storms	No. Supercells	No. AACPs	Tornado-Hail-Wind Reports	Significant Tornado-Hail-Wind Reports
10 May 2014	112	18	18	5-363-29	2-34-1
11 May 2014	330	15	32	41-398-73	8-40-32
21 May 2014	54	1	12	5-79-1	0-0-0
19 May 2015	329	35	30	36-121-18	0-33-3
24 May 2015	123	2	21	10-98-0	1-24-0
27 May 2015	387	25	32	19-392-12	1-154-1
28 March 2017	380	23	23	9-123-46	1-11-9
5 April 2017	453	13	58	27-701-331	5-54-70
16 May 2017	645	14	37	15-274-83	2-59-22
18 May 2017	535	21	54	73-331-176	3-53-52
28 June 2017	548	11	35	28-126-134	5-27-30
29 June 2017	484	10	32	1-168-98	0-29-29
19 July 2017	204	6	21	3-27-29	0-4-8
Total	**4583**	**194**	**405**	**272-3201-1030**	**28-522-257**

Given the subjective nature of case selection, some biases can be introduced. For example, all days feature many severe weather reports, possibly biasing the data towards severe occurrence. However, by separating severe and non-severe storms during these days through methods discussed in later sections, this bias can be reduced. Further discussion of potential biases resulting from case selection in Chapter 7.

2.2 Radar Observations

The NEXRAD network is comprised of more than 140 S-band (10-11 cm wavelength) Weather Surveillance Radar - 1988 Doppler (WSR-88D) systems over the CONUS (Crum and Alberty 1993). WSR-88D systems observe clouds and precipitation on a polar grid in range, azimuth, and elevation relative to the location of the radar. Each WSR-88D provides data volumes at approximately 5-min intervals when sampling convection. All NEXRAD observations used in this study were obtained at a range resolution of 250 m, an azimuthal resolution of 0.5 degrees for the lowest 3-5 elevations and 1.0 degree otherwise, and typically at 14 elevations per volume. The variables measured and used in this study include Z_{H}, the radial Doppler velocity ($\mathrm{V_R}$), Z_{DR}, differential propagation phase shift (ϕ_{DP}), and the copolar correlation coefficient (ρ_{HV}). Due to the range dependence of ϕ_{DP} measurements, we compute half the range derivative of ϕ_{DP} (the specific differential phase or $\mathrm{K_{DP}}$) for analysis. NEXRAD data were obtained from the National Centers for Environmental Information (NCEI, NOAA/NESDIS/NCEI 1991) for all of the 30 events used in this study (see Table 2.1).

NEXRAD data were merged into a large-area, four-dimensional space-time composite for analysis using the methods described in Homeyer and Bowman (2017) and references therein, referred to as GridRad. These volumes feature 5-minute temporal resolution, $\sim$2 km horizontal grid spacing, 0.5-km vertical grid spacing in the first 7 km ASL and 1-km above. Due to complexities involved in merging $\mathrm{V_R}$, its azimuthal derivative (azimuthal shear) and radial derivative (radial divergence) are instead used. The weighted

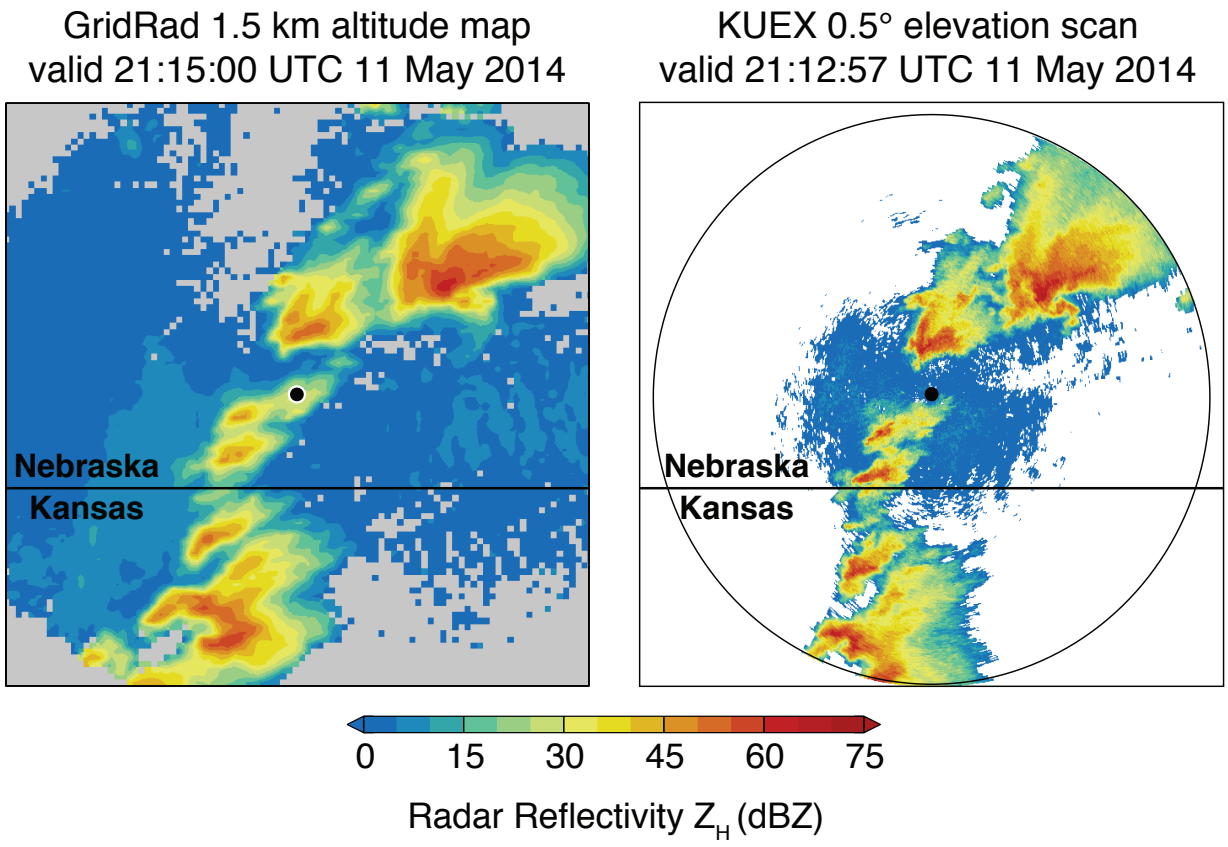

Figure 2.1: Maps of (left) GridRad radar reflectivity at a constant altitude of 1.5 km ASL and (right) radar reflectivity from the 0.5° elevation scan of the Hastings, Nebraska WSR-88D (ICAO code KUEX) for a cluster of storms occurring on 11 May 2014, including a severe hail-producing supercell to the north and east of the KUEX radar. The low-level KUEX scan is within 3 minutes of the GridRad analysis. Black dots in each map indicate the location of the KUEX radar.

space and time averaging employed in the GridRad algorithm reduces uncertainty in the radar variables, allowing for prominent and spatially consistent features to be retained. While some detail at scales near and below that of the GridRad data that is present in the native NEXRAD volumes and potentially relevant to hail identification is lost as a result of the space and time averaging, such losses are expected to have a minimal impact on the analysis. An example comparison between GridRad at 1.5 km constant altitude (ASL) and the lowest elevation scan from a single radar can be seen in Figure 2.1.

Prior to merging of the radar variables using the GridRad algorithm, several quality-control techniques are employed. First, only observations with $\rho_{\mathrm{HV}} > 0.5$ are used in order to limit the impact of non-meteorological scatterers on the resulting GridRad volumes. Second, ϕ_{DP} is smoothed using a 7.5-km radial running-mean filter to minimize noise prior to the $\mathrm{K_{DP}}$ calculation, which is calculated using first-order centered differencing. Third, Z_{DR} observations from individual radars undergo objective bias correction using a "natural scatterer" approach that targets snow aggregates at altitudes above the 0°C temperature level (e.g. Ryzhkov et al. 2005a). For more information involving these techniques and justification for their use, see the extensive review in Homeyer and Kumjian (2015). Once single radar observations are quality-controlled, further corrections are made to full GridRad volumes. Low-confidence echoes, which include those occurring either at great distances from contributing radar locations and/or within a GridRad volume sampled infrequently by the NEXRAD network are removed (Homeyer and Bowman 2017). To reduce observations resulting from non-meteorological scatterers, echoes are removed when $Z_{\mathrm{H}} < 40$ dBZ and $\rho_{\mathrm{HV}} < 0.9$.

When $\mathrm{V_R}$ exceeds the maximum detectable value at a particular operating mode (i.e., the Nyquist velocity), it becomes aliased or "folded". De-aliasing is completed prior to computing the $\mathrm{V_R}$ derivatives by implementing the open-source routine "dealias_region_based" from the Python ARM Radar Toolkit (Py-ART; Helmus and Collis 2016). Once the $\mathrm{V_R}$ fields are de-aliased, a 3x3 median filter and a 5-gate running-mean range filter are applied

to reduce random noise. The azimuthal and radial derivatives are then computed using first-order centered differencing, where the spatial resolution of the derivatives varies with range and elevation to account for changes in resolution associated with the native polar grid of the NEXRAD observations (i.e., we compute azimuthal shear at approximately equivalent spatial resolution at all ranges from the radar). All data quality measures outlined above ensure that GridRad volumes contain high-confidence observations of storms.

As denoted in Table 2.3, radar observations provide information on physical and dynamical processes taking place in individual storms. Z_{H} is proportional to the sixth moment of hydrometeor size distribution, providing information on particle size. Z_{DR} is the ratio between reflectivity factors at horizontal and vertical polarizations and provides greater detail of particle shape. ϕ_{DP} provides information on particle type and phase and its derivative, $\mathrm{K_{DP}}$, provides concentration information of non-spherical particles within the observed volume. Azimuthal shear is approximately half the total vertical vorticity (assuming circular vortices), and radial divergence is approximately half of the total horizontal divergence (though some limitations exist at high altitudes where the radar observations contain some velocity contributions from the vertical component of the wind). More information involving the physical interpretation of these radar observation can be found in several textbooks and review papers (e.g. Herzegh and Jameson 1992; Doviak and Zrnić 1993; Hubbert et al. 1998; Zrnić and Ryzhkov 1999; Straka et al. 2000; Bringi and Chandrasekar 2001; Ryzhkov et al. 2005b; Kumjian 2013a,b,c).

We focus our analysis on quantitative parameters rather than qualitative ones (such as the HCA or SL3D) as their binary nature limits the parameter space to evaluate. Numerous physical and dynamical parameters were calculated from GridRad data to quantify their relationship to hail events (Table 2.3). To avoid noise within weak echo in the polarimetric and dynamical measurements, Z_{DR}, $\mathrm{K_{DP}}$, ρ_{HV}, azimuthal shear, and radial divergence were only analyzed within echo of $Z_{\mathrm{H}} \geq 30$ dBZ. Computation of existing SP and DP radar parameters of hail detection was performed as follows. $\mathrm{H_{DR}}$, from Aydin et al. (1986) and

Table 2.3: All variables analyzed throughout hail detection analysis, categorized by their source (radar, satellite, or lightning) and representation (physical or dynamical storm characteristic).

	Radar	Satellite	Lightning
Dynamical	Half Rotation Half Divergence	mAMV Cloud Top Vorticity mAMV Cloud Top Divergence	N/A
Physical	Z_H, Z_{DR}, ρ_{HV}, H_{DR} MESH VIL Density Z_{DR} Column Altitude K_{DP} Column Altitude Echo Top Altitude	OT Probability OT IR Brightness Temperature Visible Texture Rating	Total Flash Density In-Cloud Flash Density Cloud-To-Ground Flash Density

Depue et al. (2007), combines the polarimetric variables Z_{H} and Z_{DR} to identify deviations in scattering characteristics of hydrometeors from that expected of rain using a piece-wise linear Z_{H}-Z_{DR} relationship:

$$\begin{aligned} \mathrm{H_{DR}} &= Z_{\mathrm{H}} - f(Z_{\mathrm{DR}}), \text{ where} \\ f(Z_{\mathrm{DR}}) &= 27, (Z_{\mathrm{DR}} \leq 0 \text{ dB}), \\ &= 19\, Z_{\mathrm{DR}} + 27, (0 \leq Z_{\mathrm{DR}} \leq 1.74 \text{ dB}), \\ &= 60, (Z_{\mathrm{DR}} > 1.74 \text{ dB}). \end{aligned} \tag{2.1}$$

MESH, from Witt et al. (1998a), is a commonly used metric to estimate hail occurrence and size. It is based on a power law relationship to the severe hail index (SHI), which is defined as:

$$
\begin{aligned}
W(Z) &= 0, (Z \leq Z_L), \\
&= \frac{Z - Z_L}{Z_U - Z_L}, (Z_L < Z < Z_U), \\
&= 1, (Z \geq Z_U)
\end{aligned}
\tag{2.2}
$$

$$
\begin{aligned}
W_T(H) &= 0, (H \leq H_0), \\
&= \frac{H - H_0}{H_{m20} - H_0}, (H_0 < H < H_{m20}), \\
&= 1, (H \geq H_{m20})
\end{aligned}
\tag{2.3}
$$

$$
\dot{E} = 5\ x\ 10^{-6}\ x\ W(Z) 10^{0.084Z} \tag{2.4}
$$

$$
SHI = 0.1 \int_{H_0}^{H_T} W_T(H) \dot{E}\ dH. \tag{2.5}
$$

where Z is Z_{H}, Z_L is 40 dBZ, Z_U is 50 dBZ, H is the height above ground level (AGL) of the radar observation, H_0 is the collocated height AGL of the 0°C level, and H_{m20} is the height AGL of the -20°C level. A simple power law relationship between SHI and MESH was empirically developed to fit the 75$^{\mathrm{th}}$ percentile of approximately 150 hail reports, resulting in the commonly used relation:

$$
MESH = 2.54 (SHI)^{0.5}. \tag{2.6}
$$

VIL density, from Amburn and Wolf (1997), is VIL divided by the Z_{H} = 18.5 dBZ echo top height, where VIL is defined as:

$$
VIL = \sum\ 3.44\ x\ 10^{-6} [(Z_i\ +\ Z_{i+1})/2]^{4/7} \Delta h. \tag{2.7}
$$

where Z_i and Z_{i+1} are Z_{H} at the bottom and top of a layer (in mm^6 m-3) and Δh is the layer depth (in m). While most prior studies compute VIL density by dividing VIL by the echo top height, this study computes VIL density by dividing by the echo depth (the sum

of radar-observed layers where $Z_H \geq 18.5$ dBZ). This approach helps to normalize all radar observations by their three-dimensional coverage. For large regions observed by NEXRAD radars, measurements are not available at lower altitudes and gaps in radar coverage may exist aloft. Thus, VIL density calculated using only the echo top can be biased low in cases where coverage is limited or reduced.

While MESH and VIL density vertically integrate Z_H to estimate the amount of large particles in a volume, other methods use physical parameters at one level to infer updraft characteristics and hail potential. For example, liquid hydrometeors can be lofted above the environmental freezing level in strong convective updrafts, in which case they are referred to as supercooled drops. These convectively lofted supercooled drops are observed in DP radar observations as large positive values of Z_{DR} and K_{DP} above the freezing level. When vertically continuous in altitude, these features are referred to as Z_{DR} and K_{DP} columns and have been argued to be of potential use in identifying strong convective updrafts prior to production of severe weather (Kumjian and Ryzhkov 2008; Kumjian et al. 2014; Snyder et al. 2015). We identify Z_{DR} or K_{DP} columns for use in this study as the maximum height where $Z_{DR} \geq 1.5$ dB or $K_{DP} \geq 1.5$ deg km^{-1}. Echo top heights, the maximum altitude where Z_H exceeds a specified threshold value, are also correlated with the intensity of upward motion within a storm and were computed for analysis in this study. While echo top heights for multiple thresholds were analyzed (Z_H = 10, 20, 30, and 40 dBZ), only results from 40 dBZ echo top altitudes are discussed in this paper. Figure 2.2 shows example GridRad maps of many of the variables analyzed in this study.

Supercell storms were identified using GridRad observations for all severe weather days via a combination of objective and subjective methods (Sandmæl 2017). First, potential supercell storms were objectively identified by searching for storm tracks with long duration ($\geq$ 60 min), strong azimuthal shear (or rotation, $> 3.5x10^{-3}s^{-1}$, which helps to discriminate between non-rotating or weakly rotating storms and supercells), and tall 40 dBZ echo tops ($\geq$ 12 km during any point in storm lifecycle). Objectively selected candidate storms

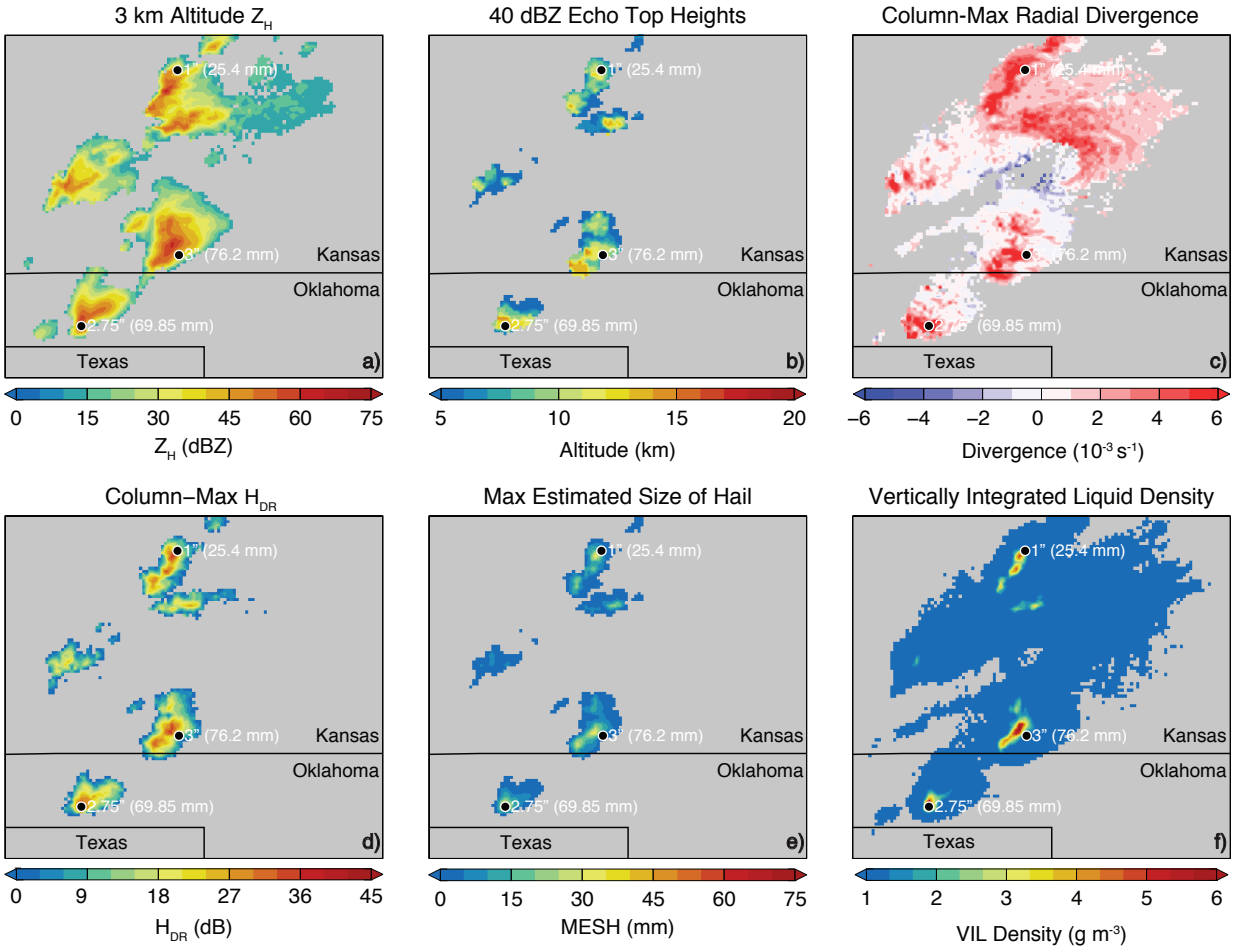

Figure 2.2: GridRad maps of (a) radar reflectivity at an altitude of 3 km ASL, (b) 40 dBZ echo top heights, (c) column-maximum radial divergence, (d) column-maximum hail differential reflectivity, (e) maximum estimated size of hail, and (f) vertically integrated liquid density, valid at 20:50 UTC on 16 May 2017. Black circles in each map indicate hail reports valid at this time, with corresponding hail sizes noted in white text.

were then subjectively evaluated to confirm or deny supercellular characteristics. In particular, in order to classify a storm as a supercell, the following conditions were sought: 1) obvious deviant motion relative to neighboring storms, 2) a hook echo at altitudes ≤ 3 km, 3) a bounded weak echo region apparent in upper-level maps or vertical sections of radar reflectivity, or 4) polarimetric signatures such as the differential reflectivity arc at low levels (Fujita 1958; Browning and Donaldson 1963; Lemon and Doswell 1979; Doswell and Burgess 1993; Kumjian and Ryzhkov 2008). If at least two of these conditions were met, the storm was categorized as a supercell. Note that this method may miss some weak supercells as well as supercells embedded within organized convective systems, but all tornadic storms and AACP-producing storms (regardless of whether or not they passed the initial objective duration, shear, and tall echo top identification) were manually evaluated to confirm supercell or non-supercell characteristics.

2.3 Satellite Observations

GOES imagery was retrieved from both NCEI (NOAA/NCEI 2013) and the University of Wisconsin-Madison Space Science and Engineering Center (University of Wisconsin - Madison Space Science and Engineering Center 2011). Data from GOES–13,–14, and –16 were used for analysis in this study. Beginning in March 2017, preliminary data from GOES–16 were available until the system was officially declared operational in December 2017. While GOES-16 can operate in 30-sec imagery mode, 1-minute temporal resolution was the finest that was analyzed in this study. For 17 of 30 overall severe weather days (bolded in Table 2.1 and all cases in Table 2.2), 1-minute imagery was available for analysis through either GOES–14 SRSOR mode or GOES–16 Flex Mode. GOES–13 was used for the remaining days, where temporal resolution ranged from 5- to 15-minutes.

GOES–13 and –14 data include 1-km pixel^{-1} spatial resolution for imagery from the 0.65-μm visible channel and 4-km pixel^{-1} spatial resolution for imagery from the 10.7-μm

IR channel. GOES–16 data includes 0.5-km pixel^{-1} spatial resolution for 0.65-μm visible channel imagery and 2-km pixel^{-1} spatial resolution for 10.3-μm IR channel imagery (Schmit et al. 2005, 2014). For 29-30 June 2017, 10.3-μm channel data were unavailable, so the 11.2-μm channel was used as a substitute. The resulting satellite parameters were not biased by this difference, as the 11.2-μm channel is only slightly more sensitive to water vapor.

Using these GOES data, products based on both visible and IR imagery were derived for analysis, including texture and temperature characteristics of OTs and mesoscale atmospheric motion vectors (mAMVs). OTs appear in visible imagery when strong convective updrafts rise above the level of neutral buoyancy, exhibiting more turbulent texture than the surrounding anvil. For severe storms, the level of neutral buoyancy is often the tropopause and OTs that extend to higher altitudes within the stratosphere are indicative of strong upward motion. Bedka and Khlopenkov (2016) developed an objective OT identification/probability and a product quantifying OT texture in visible imagery (named "visible texture rating"), which we employ for analysis in this study. Dynamical products based on mAMVs were derived through tracking of satellite features in 1-min visible and IR imagery via the algorithms outlined in Bedka and Mecikalski (2005), Bedka et al. (2009), and Apke et al. (2016). Two variables, cloud top vorticity (CTV) and cloud top divergence (CTD), are retained for analysis. Since the mAMV algorithms require 1-minute temporal resolution, analyses of CTV and CTD are not possible for cases using GOES–13 data. In addition, mAMV products are also unavailable for cases using GOES–16 data since the algorithms have not yet been adapted to the finer spatial resolution of GOES-16 imagery. Therefore, CTV and CTD are only available for cases with GOES–14 data (denoted by "*" in Table 2.1).

2.4 Environmental Data

We use numerical model analyses to obtain atmospheric quantities needed to calculate certain parameters (i.e. MESH). Hourly analyses from the Rapid Refresh (RAP) numerical weather model operated by the National Centers for Environmental Prediction (NCEP) were used (NOAA/NCEP/ESRL 2012). Analyses are available at 13-km horizontal resolution and 50 vertical levels. Environmental parameters used for analysis include the tropopause, and T = 0 °C, −10 °C, and −20 °C altitudes.

2.5 Lightning Data

Lightning data from the Earth Networks Total Lightning Network (ENTLN) were included given its utility to identify strong, mixed-phased updrafts as discussed in Chapter 1. The network detects lightning using pulses in vertical electric field measurements from parts of the 1 Hz to 12 MHz frequency range from over 700 sites across the CONUS (Liu and Heckman 2010). Individual pulses are located in space and time by statistically solving over-determined electrical signal time-of-arrival equations using measurements from at least 5 stations. Sources close together in space and time are grouped into flashes (consisting of both cloud-to-ground and intra-cloud flashes), which were binned into $0.08^\circ \times 0.08^\circ$ longitude-latitude flash extent density grids for analysis in this study (an area of $\sim$64 km^2) to emulate what could be provided by the GOES-16 Geostationary Lightning Mapper (Goodman et al. 2013) once the data becomes fully operational. Final analysis of the flash extent density grids was conducted using the maximum of the total lightning flash density along each storm track.

2.6 National Weather Service Warnings

NWS severe thunderstorm and tornado warnings were only needed for the second main objective, and were retrieved through an online archive, hosted by the Iowa Environmental

Mesonet (IEM), which is a volunteer effort based out of Iowa State University (Iowa State University 2018). The database includes the time of issuance, time of expiration, and coordinates of the polygon for each severe thunderstorm and tornado warning issued by the NWS. Warnings were linked with a storm if the storm passed through the polygon during valid warning times (i.e. any time between issuance and expiration). Most warnings also featured written text that specified certain hazard magnitudes expected during the warning time period. These hazard "tags" can include maximum wind gust and/or hail size, however, only expected hail size information was used for analysis.

2.7 Severe Weather Reports

The NOAA National Centers for Environmental Information (NCEI) hosts the severe weather data inventory (SWDI) storm event database that contains the time, duration (if available), location, magnitude and source of all confirmed U.S. severe weather reports (hail reports $\geq$ 0.75 inches in diameter), which we use for analysis of the observational data in this study (NOAA/NCEI 2014). We also utilize reports from the NOAA Storm Prediction Center (SPC) (NOAA/NWS/SPC 1955) for the objective hail analysis only. SPC reports also provide the time, location, and size of severe hail reports, which is defined as those $\geq$ 1 inch in diameter (all cases occurred after the severe hail criteria change from $\geq$ 0.75 inch). Severe weather is classified as hail reports $\geq$ 1 in, wind reports $\geq$ 50 knots, and/or confirmed tornado reports, and significant severe weather includes hail reports $\geq$ 2 in, wind reports $\geq$ 65 knots, and/or tornadoes resulting in $\geq$ EF2 damage. While these reporting databases provide the most complete account of trusted/confirmed hail reports in the CONUS, there are many well-established reporting biases that pose unique challenges to analysis and verification (see discussion in Section 1).

2.8 Storm Track Development

All datasets outlined in the previous subsections were analyzed on an individual storm (i.e. "cell") basis in this study. Storm tracks were developed using objective radar echo top tracking based on the methods outlined in Homeyer et al. (2017) and Sandmæl (2017). Briefly, storm tracking is accomplished through time-linking of 40 dBZ echo top maxima within echoes classified as convection by the Storm Labeling in 3 Dimensions (SL3D) algorithm (Starzec et al. 2017). Objectively tracked storms were reviewed for each case to manually identify and merge discontinuous tracks that correspond to the same storm. These quality-controlled tracks were linearly interpolated in time to obtain 1-minute temporal resolution in order to match the finest resolution of the available datasets. Observations and environmental parameters were then extracted along the tracks within a 10-km radius of the storm location for analysis. Finally, reports were linked to the closest storm within a maximum allowed distance of 10 km from the report location.

Storm tracks were then separated into two groups: the "hail" population and the "no-hail" population. The hail population consisted of storms that had at least one hail report during its lifetime. The no-hail population has lower confidence than its hail counterpart, as one cannot definitively confirm a lack of hail occurrence (i.e., hail may have occurred but was simply not reported). To account for population-driven biases of the no-hail group, storms that occurred in regions of low population were excluded from the analysis. Population density from the Center for International Earth Science Information Network for 2015, the most recent available year, was extracted along each storm track. The distance between storm location and the nearest city was calculated and similarly linked to the storm tracks, but the results using population density were clearer. Requiring a lower limit of population density or an upper limit of distance to nearest city aims to retain no-hail storms (or portions of such storms) for analysis with greater confidence (i.e., a sufficient number of potential witnesses are in the path of the storm, such that severe hail is more likely to be reported by at least one individual if it occurred). By comparing cumulative frequency distributions

of radar observations (each representing a potential hail event) and hail reports, both as a function of population density, it is possible to identify if/when the relationship between hail report frequency and population density is not one-to-one (or at least linear). Figure 2.3 provides this comparison for all observations analyzed in this study. This comparison reveals that hail reports and population density are linearly correlated below a population density threshold of approximately 25 people per square mile. Beyond this threshold, hail reporting and population density are no longer linearly correlated and therefore expected to be less biased. Based on this evaluation, we imposed a lower limit of 25 people per square mile population density to the no-hail population for further analysis, minimizing the analyzed observations to segments of no-hail storms with population density higher than this value. The resulting sample of no-hail storm observations retained was approximately 14% of the original sample.

When linking AACPs to storm tracks, GridRad track data were used to identify the storm cell ID number corresponding to the AACP-producing updraft. Satellite imagery of convection can be complex, especially when rapid storm growth and decay of new cells occur in close proximity. Therefore, it is sometimes difficult to ensure that one is monitoring the same cell in a sequence of images. Understanding severe weather characteristics before, during, and after an AACP is embedded in the second main objective, so ensuring that we correctly associate an AACP with its parent updraft is critical. Pairing the GridRad storm tracks with GOES data reduces uncertainty. GridRad storm tracks were corrected for parallax based on the 10 dBZ echo top height for co-location with GOES data. AACP lifetime was determined by identifying sustained emission of warm or textured cloud material from an OT region along the same GridRad storm track. An AACP must be produced continuously for 10+ mins to be considered in this analysis. GOES data must be available throughout the lifetime of the storm for an AACP to be included in AACP analyses, though GOES image outages of up to five minutes were accepted. Occurrences

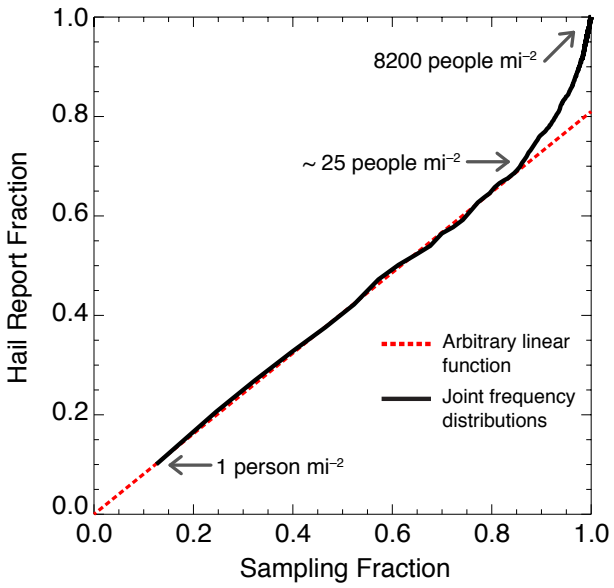

Figure 2.3: Comparison of cumulative frequency distributions as a function of population density for all radar observations (sampling) and hail reports analyzed in this study. The thick black line is the joint cumulative frequency distribution relationship and the red line is an arbitrary linear function to demonstrate deviation of the observed relationship from linearity. Values of population density along the joint cumulative frequency distribution are labeled at several points within the plot.

of an AACP-producing cell changing ID number due to brief cell decay, splitting, or other complex reasons were noted.

Most cells continuously produce an AACP for a period of time and then decay or, in the case of a supercell, split into two cells that may each produce their own AACP. In some situations, a cell will produce an AACP for some time, then production stops while a strong radar echo persists. The cell can then restart plume production at some later time. All time periods of active plume production are noted in these instances. In some situations, most notably in MCSs where updraft cores are close to one another and rapidly evolving, cell ID's may change more frequently than would otherwise be expected. AACP production can be brief in these instances and challenging to link to an individual GridRad storm track. It is typically clear when plume production stops. In particular, an OT region disappears and the warm, textured AACP detaches from the OT and advects downstream. An AACP and associated warm anomaly (if present) can be evident for several hours after production ceases, and the AACP eventually sublimates (Smith et al. 2016; Homeyer et al. 2017).

Chapter 3

Hail Detection Methods

3.1 Performance Evaluation

Storms were labeled as hail-producers when a given parameter threshold was met; naturally occurring values for each parameter were tested. The ability to distinguish between hail- and non-hail-producing storms was then accessed by calculating the probability of detection (POD), false alarm rate (FAR), and critical success index (CSI). The forecast statistics used in this study are as follows:

$$\text{POD} = \frac{\text{No. correctly identified hail storms}}{\text{No. hail storms}}, \tag{3.1}$$

$$\text{FAR} = \frac{\text{No.storms misidentified as hail storms}}{\text{No. identified storms}}, \tag{3.2}$$

$$\text{CSI} = \left(\frac{1}{1-\text{FAR}} + \frac{1}{\text{POD}} - 1\right)^{-1}. \tag{3.3}$$

Additionally, parameters were combined into multi-variable indicators such that all parameters must exceed their corresponding threshold to indicate a hail-producing storm.

3.2 Statistical Significance

Statistical significance was evaluated using the two-sample Kolmogorov–Smirnov (KS) test that assessed whether the two samples originate from the same distribution. This was done using the following relations:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,x]}(X_i). \tag{3.4}$$

$$D_n = \sup_x |F_n(x) - F(x)|. \tag{3.5}$$

where $F_n(x)$ is the cumulative distribution function calculated for each sample and D_n is the critical value. The hull hypothesis (that the two samples did originate from the same distribution) is rejected when the critical value is less than the desired confidence interval, α. This study used a confidence interval of α = 0.01 (confidence of 99%, such that distributions were significantly different if $D_n > 0.01$).

Chapter 4

AACP-Severe Weather Linkage Methods

4.1 AACP Identification

GOES visible and IR imagery were animated and analyzed using the McIDAS-V software package (Achtor et al. 2008) to determine the starting and ending time of AACP production by an individual storm updraft. AACP identification was based on IR imagery and supplemented by visible imagery when available. The identification process was subjective, driven by the extensive image analysis experience described by Setvák et al. (2013); Homeyer (2014); Bedka et al. (2015); Homeyer et al. (2017). Areas of warm IR BTs emanating from an OT are identified, are often 10 K warmer than the surrounding anvil, and appear anomalous relative to non-AACP storms that continuously produce cold anvil cloud. Animated visible imagery was used during the daytime to determine if the warm anomaly was associated with visible shadowing and/or a variation in texture typical of an AACP. Over 70% of AACPs were generated for less than one hour but eight storms (2% of the population) produced a continuous AACP for more than four hours. The distribution of AACP duration is similar to the enhanced-V events analyzed by McCann (1983). Figure 4.1 gives three examples of AACP signatures in visible and IR imagery that were present on 19 July 2017.

As discussed above, it is important to note that an AACP is not always warm. Heymsfield et al. (1983) hypothesized that warming adjacent to the OT was generated by subsidence along the descending branch a breaking gravity wave, which he called the "close-in warm area". Fujita (1974) referred to this area as the "warm trench". Hypotheses for cold AACPs include 1) sedimentation of large ice crystals that reduces the cloud optical depth, allowing colder tropospheric anvil beneath the AACP to dominate the radiative signal, 2)

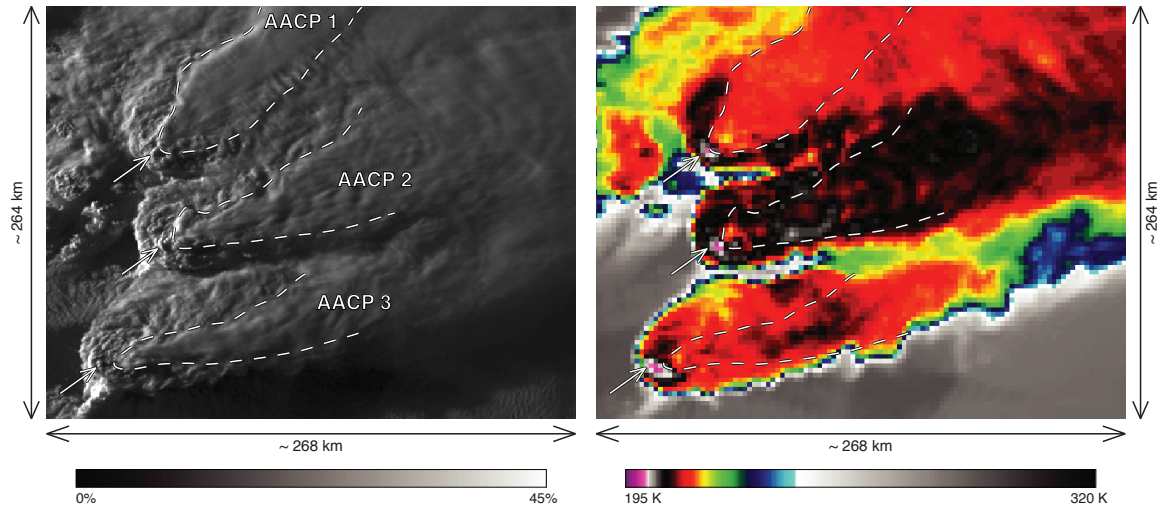

Figure 4.1: Three examples of AACPs in GOES-16 0.64-μm VIS albedo (left) and 10.4-μm IR brightness temperature (right). AACPs are enclosed in white long-dashed lines. White arrows point to overshooting tops that correspond to cold anomalies seen in IR brightness temperature (left).

AACP injection into nearly isothermal upper troposphere lower stratosphere (UTLS) environments or above-anvil layers that are cooling with height, 3) plume subsidence into layers with colder temperature, and/or 4) cooling of the local UTLS temperature through AACP sublimation. From our experience, a storm rarely produces only a cold AACP. That is, an AACP is warm in some portion of the plume, especially adjacent to the OT, or at some point during the storm lifetime. At night, there is increased uncertainty in AACP duration of unknown magnitude given the absence of visible imagery. An AACP is identified at night in this study by identifying a warm area adjacent to the OT and a narrow corridor of "plume-shaped" outflow that is continuously emitted from the OT region. AACP production ends when the OT dissipates, and the outflow channel is no longer being actively emitted. Brunner et al. (2007) suggests that a large temperature difference between the OT and AACP warm area (16 K on average), which they define as the "anvil thermal couplet", is a strong indicator of a severe storm. These complex anvil BT patterns within AACP storms illustrate the many challenges associated with physical interpretation of such

AACP BT features and the benefits of rapid-updating, multispectral imagery for plume identification.

4.2 Analysis Techniques

A total of 405 AACP-producing storms were tracked across the 13 events in addition to 4178 other non-AACP storms. GOES- and GridRad-derived characteristics of AACP storms before, during, and after AACP production are extracted for analysis. Examples of AACP and non-AACP storm tracks and severe weather reports for 4 of the 13 events are shown in Figure 6.1. Severe weather frequency and type throughout AACP storm lifetimes is also evaluated and contrasted with non-AACP storms. Relationships between AACPs and supercells are also derived to determine if the AACP can be used to identify a supercell and to contrast supercells with and without AACPs. Lastly, the relationship between AACPs, NWS warnings, and significant hail is explored to determine if an AACP detection could increase warning confidence that significant hail will occur.

Chapter 5

Hail Detection Results

5.1 Objective hail event identification

In order to evaluate the potential for hail identification using the datasets available to this study, metrics detailed in Table 2.3 within 5 min of a hail report (to mitigate the effects of report timing biases) were divided into size categories and compared to metrics at all instances of the filtered-no-hail population (see Table 5.1 for a list of categories and number of observations). Several variables were able to identify hail events well, with the overall best separation between hail and no-hail observations found in storm-maximum MESH, VIL density, and H_{DR} above and below the melting level (AML and BML, respectively, Figure 5.1). For these four variables, the no-hail distribution was significantly different from all 3 hail distribution categories, and each hail category was significantly different from one another. Joint frequency distributions of multiple parameters were also evaluated, but none improved upon the separations between hail and no-hail populations evident in the single-variable analysis.

Table 5.1: Resulting storm data post-population filter used for severe hail fall and hail storm evaluation. Distributions for reports < 1 in are not shown.

Category	No. unique observations	No. hail reports
Reports < 1 in	519	554
Reports ≥ 1 in and < 2 in	3,919	4,405
Reports ≥ 2 in	835	995
All filtered no-hail population	45,589	N/A

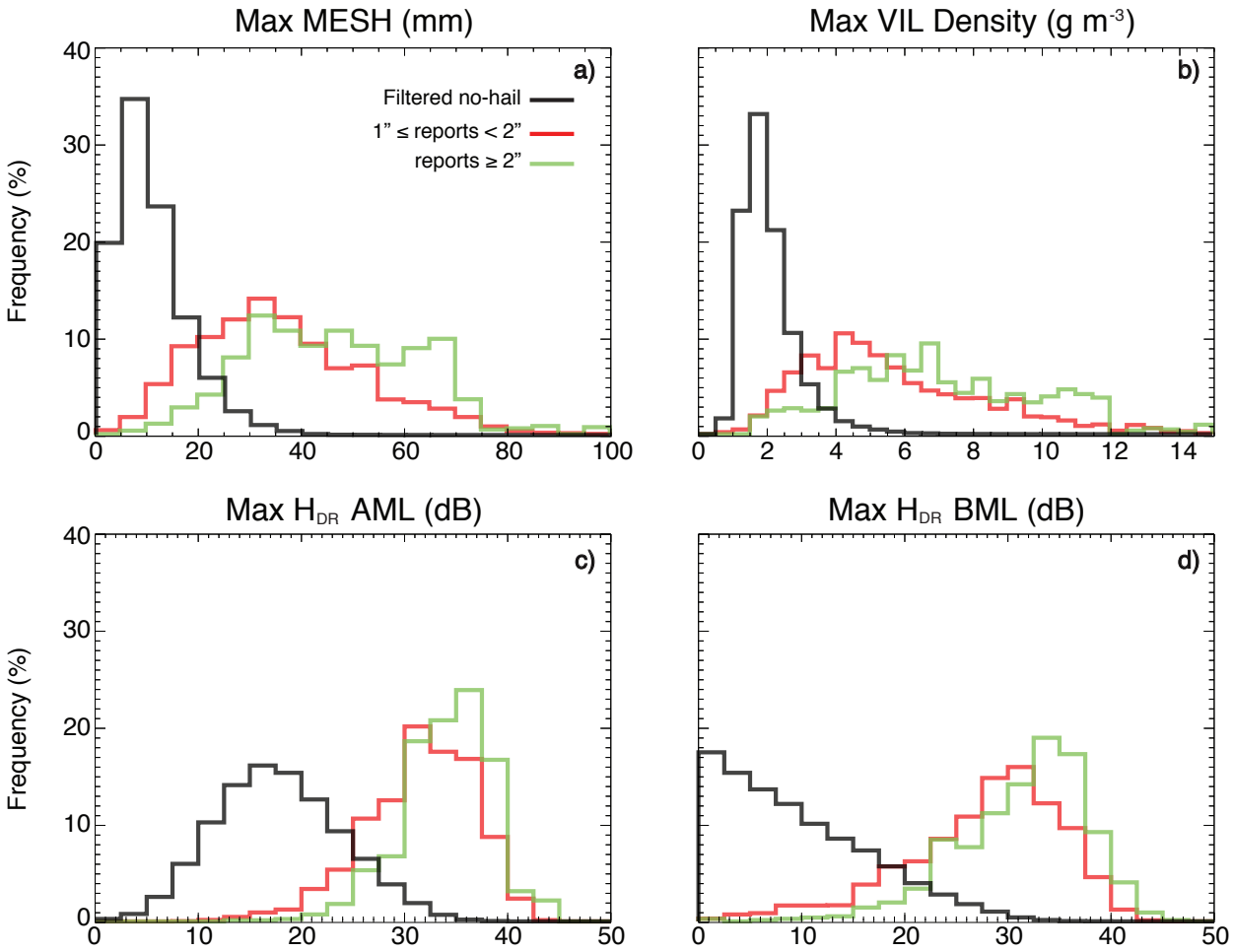

Figure 5.1: Frequency distributions of storm-maximum a) MESH, b) VIL density, c) H_{DR} above the environmental melting (0 °C) level, and d) H_{DR} below the environmental melting level within ±5 min of a hail report (colored curves) and within population-filtered segments of no-hail storms (black curves).

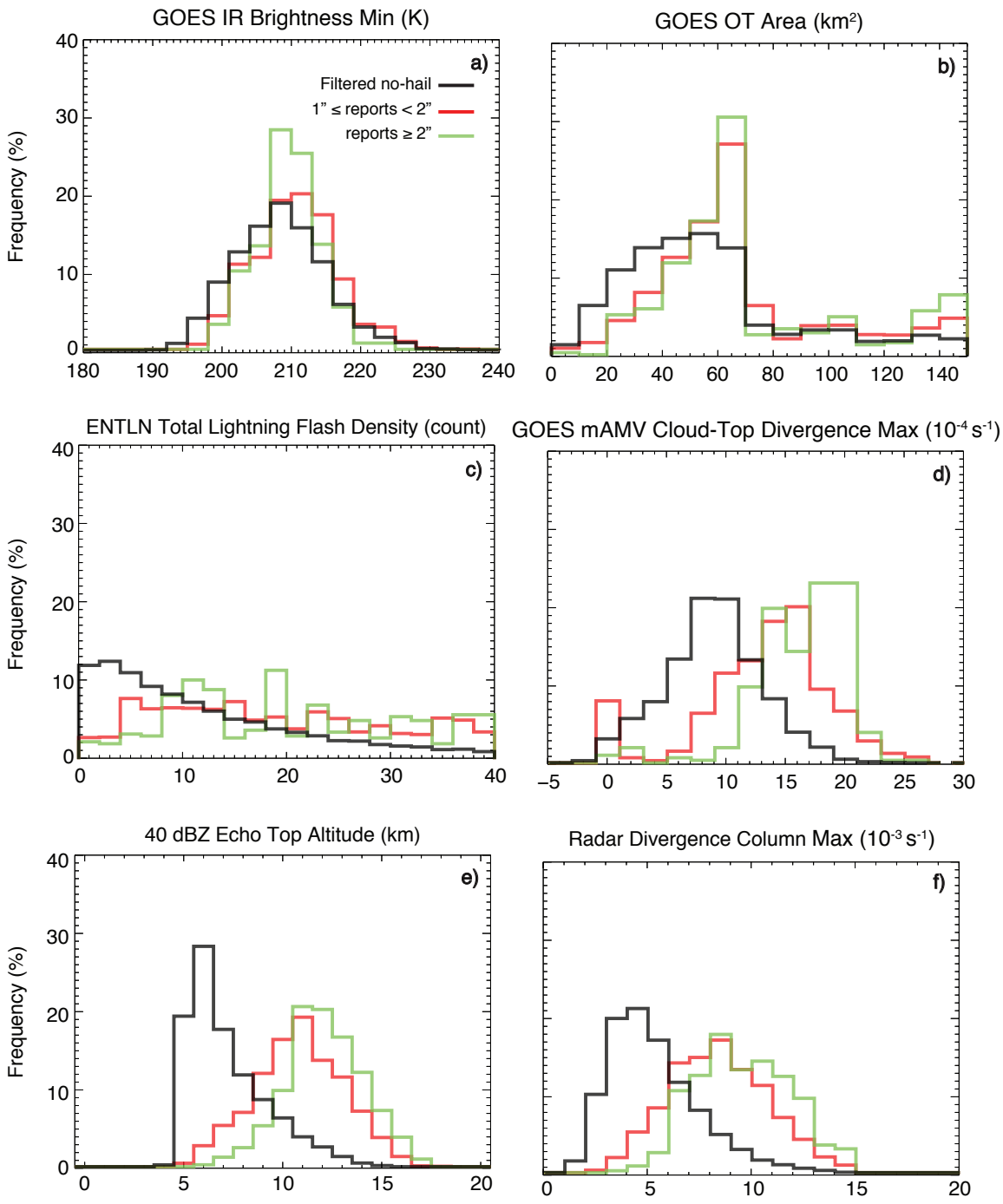

Figure 5.2: As in Fig. 5.1, but for a) storm-minimum GOES IR brightness temperature, b) GOES overshooting top area, c) ENTLN total lightning flash density, d) storm-maximum GOES mAMV cloud-top divergence, e) 40 dBZ echo top altitude, and f) storm-maximum radar divergence.

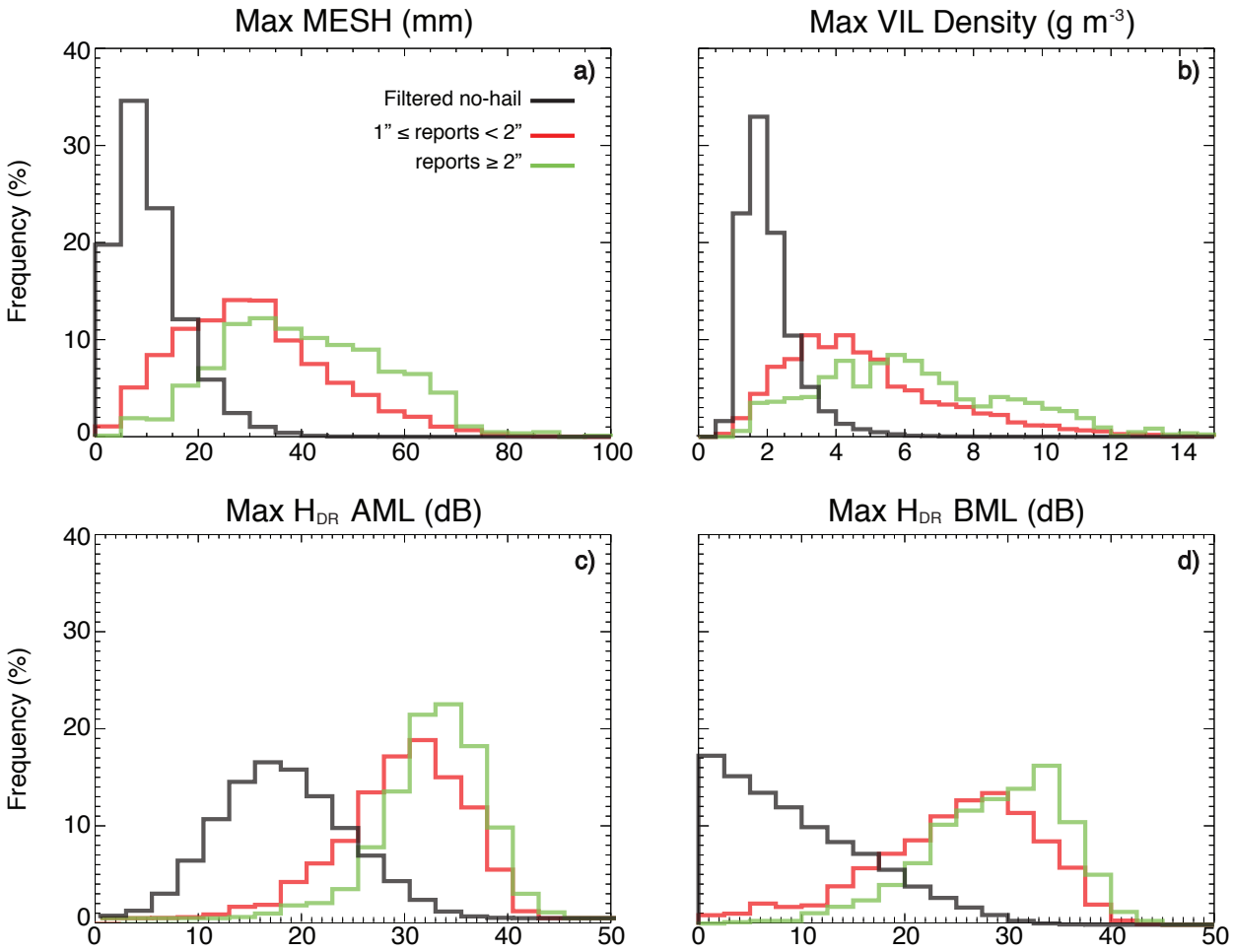

Figure 5.3: As in Fig. 5.1, but at the time of hail reports.

In comparison, several variables that have shown some utility at discriminating between severe and non-severe storms (or hail and non-hail storms) in previous studies did not show significant separation between hail and filtered-no hail populations near the time of hail events. Satellite and lightning parameters generally associated with severe weather, such as those in Figure 5.2, displayed overlap between all distributions. There is some potential for the maximum in mAMV cloud-top divergence, as it displayed the best separation of all satellite products. However, mAMVs are not routinely available or always possible due to their reliance on 1-min imagery and, as a result, these distributions only included data from 8 of the 30 severe weather days analyzed here (denoted by "*" in Table 2.1). Alternative radar parameters such as 40 dBZ echo top altitudes and radial divergence (Figs. 5.2e and

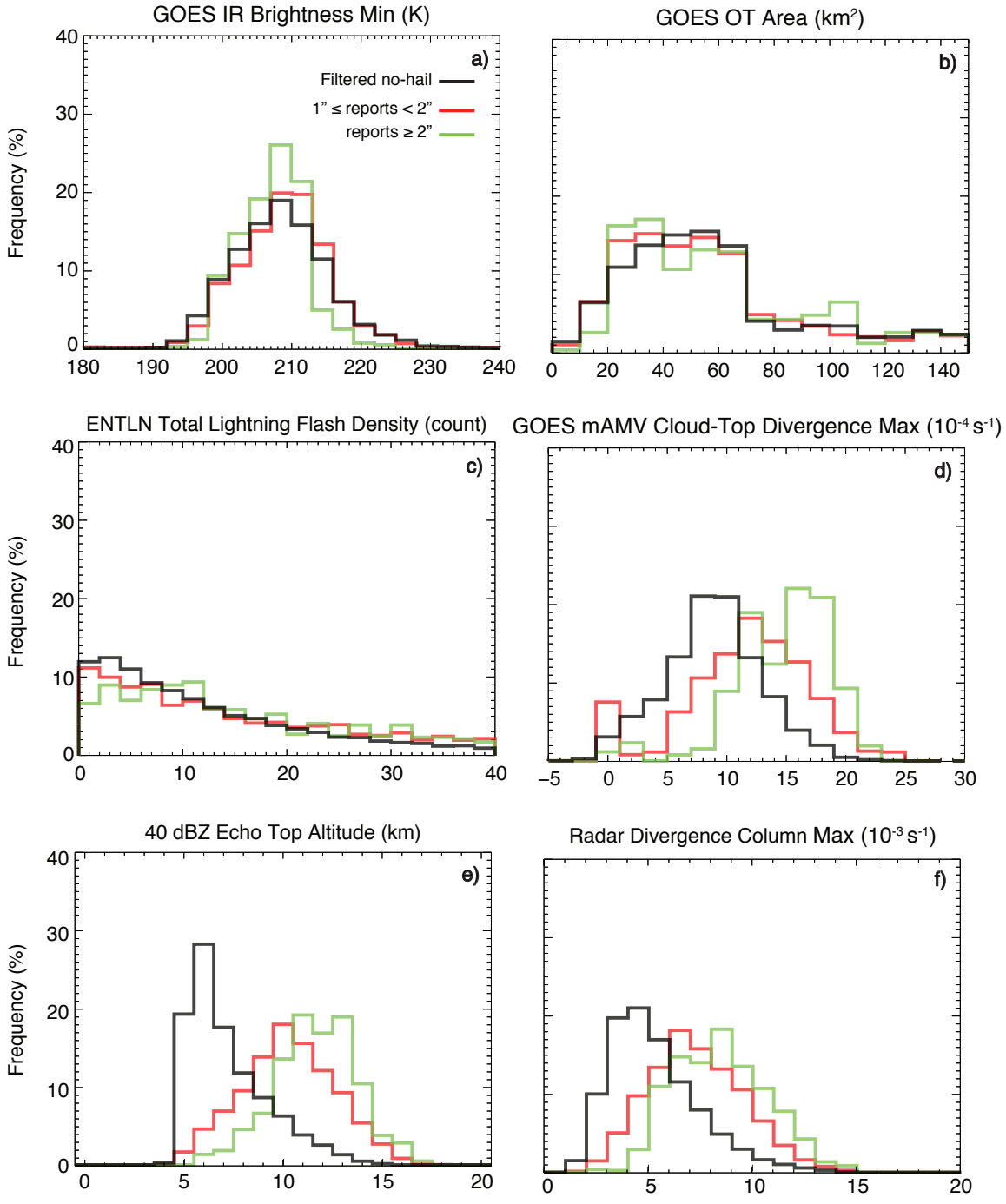

Figure 5.4: As in Fig. 5.2, but at the time of hail reports.

5.2f), were not as skillful as other radar variables previously mentioned. These results indicate that, while many products may be useful in specific severe weather scenarios, objective hail identification is best accomplished using a few select SP and DP radar parameters.

Similar comparisons were completed at the exact time of reports to determine whether or not the previous findings were a result of considering extremes in these metrics over a 10 min period (Figs. 5.3 & 5.4). These comparisons show similar behavior to the previous analysis, but with more overlap between the hail and filtered no-hail populations. Despite the increased overlap, consistent statistical significance is found for each variable. Given the consistent results between these two approaches and the increased separation between categories when allowing for $\pm$5 minutes of time error, the remaining analyses based on hail events use the $\pm$5 minute extremes of the metrics for reference.

5.2 Objective severe hail-producing storm identification

While we found that several parameters are skillful at identifying hail occurrence, verification for any objective hail identifier still remains a challenge due to the numerous known reporting biases. In order to fully assess the utility of objective identifiers, we must approach hail identification from an individual report perspective and a storm perspective. Focusing on identifying hail-producing storms rather than events allows us to better evaluate the performance of these metrics and build confidence in the results from Section 5.1. For this analysis, we are only focusing on severe hail storms ($\geq$ 1 in report), resulting in 805 hail storms. When establishing the no-hail storm population, we only evaluate storms when they are within regions of 25 people per square mile, which resulted in 3,823 no-hail storms.

The three best discriminators for hail occurrence from the previous analysis (MESH, VIL density, and H_{DR} BML) are used to determine their ability to discriminate between severe hail-producing and non-hail-producing storms. POD, FAR, and CSI for a range of

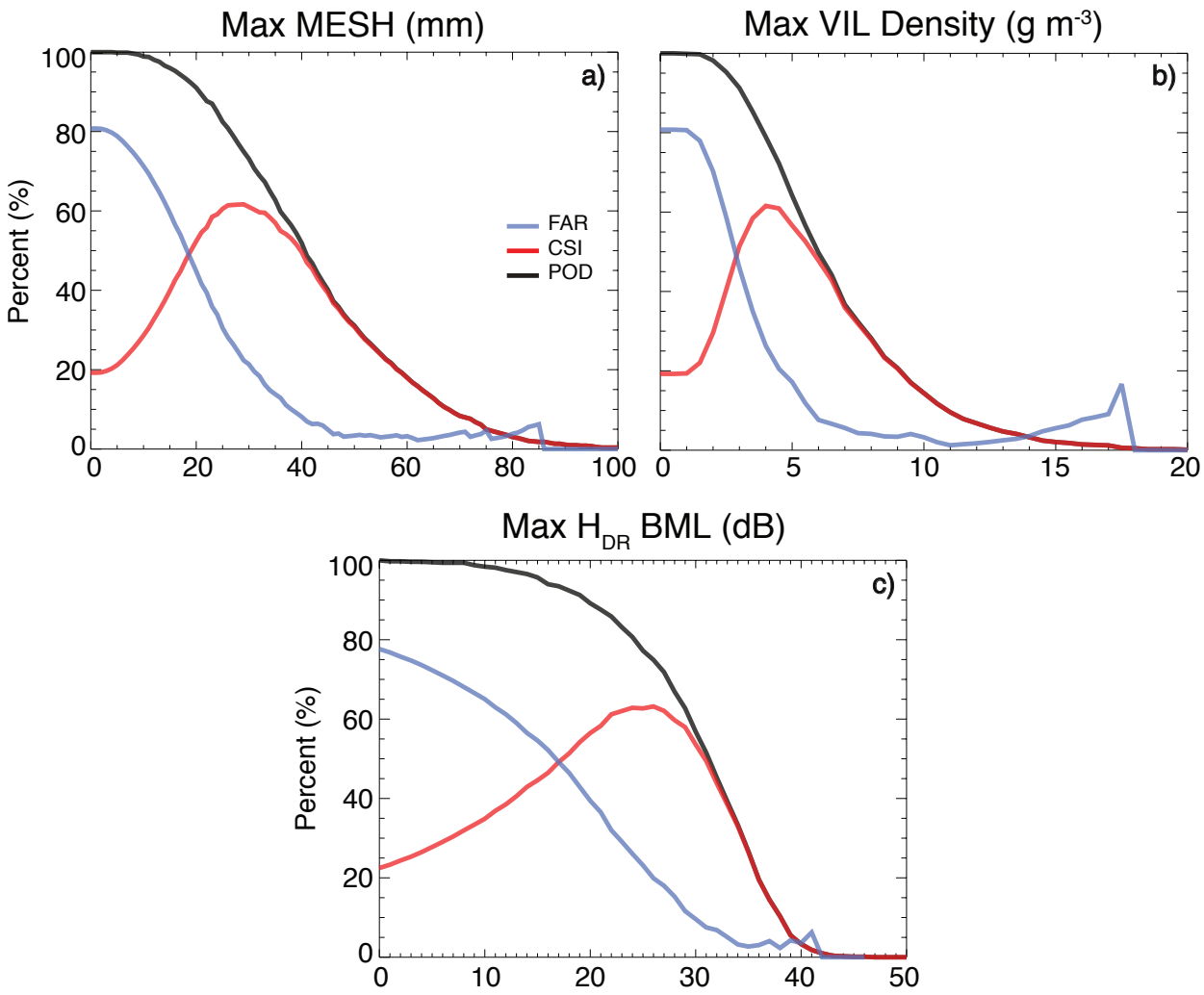

Figure 5.5: The POD, FAR, and CSI of threshold-based hail-storm identification using storm-maximum (a) MESH, (b) VIL density, and (c) maximum H_{DR} below the environmental melting level.

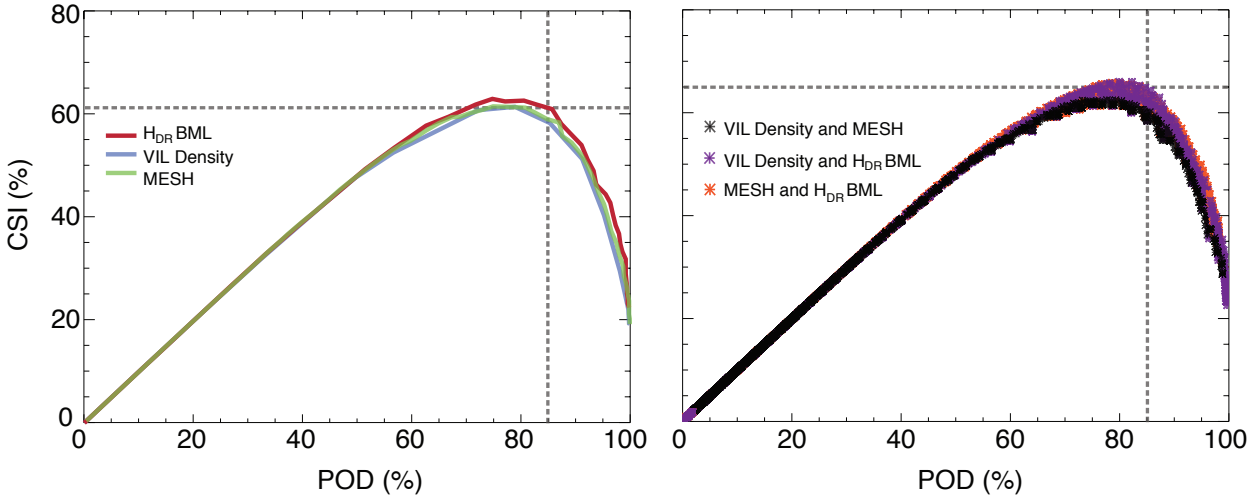

Figure 5.6: CSI as a function of POD for the range of threshold values covered in Fig. 5.5 for (left) H_{DR} below the environmental melting level (red), VIL density (blue), and MESH (green), and (right) multi-variable hail-storm identification using VIL density and MESH (black), VIL density and H_{DR} below the environmental melting level (purple), and MESH and H_{DR} below the environmental melting level (red). Lines in the single variable plots correspond to performance over the range of values in Figure 5.5, while symbols in the multi-variable plots correspond to all possible combinations of values used in the single variable performance. Vertical dashed lines in each panel indicate a POD of 85% and horizontal dashed lines indicate the highest corresponding CSI.

MESH, VIL density, and H_{DR} BML values are shown to identify the most skillful objective thresholds for each (Figure 5.5). Each variable resulted in similar performance, with a slightly broader peak and higher maximum CSI for H_{DR} BML. CSI as a function of POD better illustrates these differences (left panel, Figure 5.6). While the difference in skill between H_{DR} BML and the SP variables is statistically significant, it does not necessarily translate to practical utility as the peak CSI is only higher by approximately 2 percentage points. However, larger increases in skill can be attained when using a multiple variable approach, such that two parameters must meet the specified criteria (resulting in a scatterplot distribution between POD and CSI). Namely, using either SP variable in combination

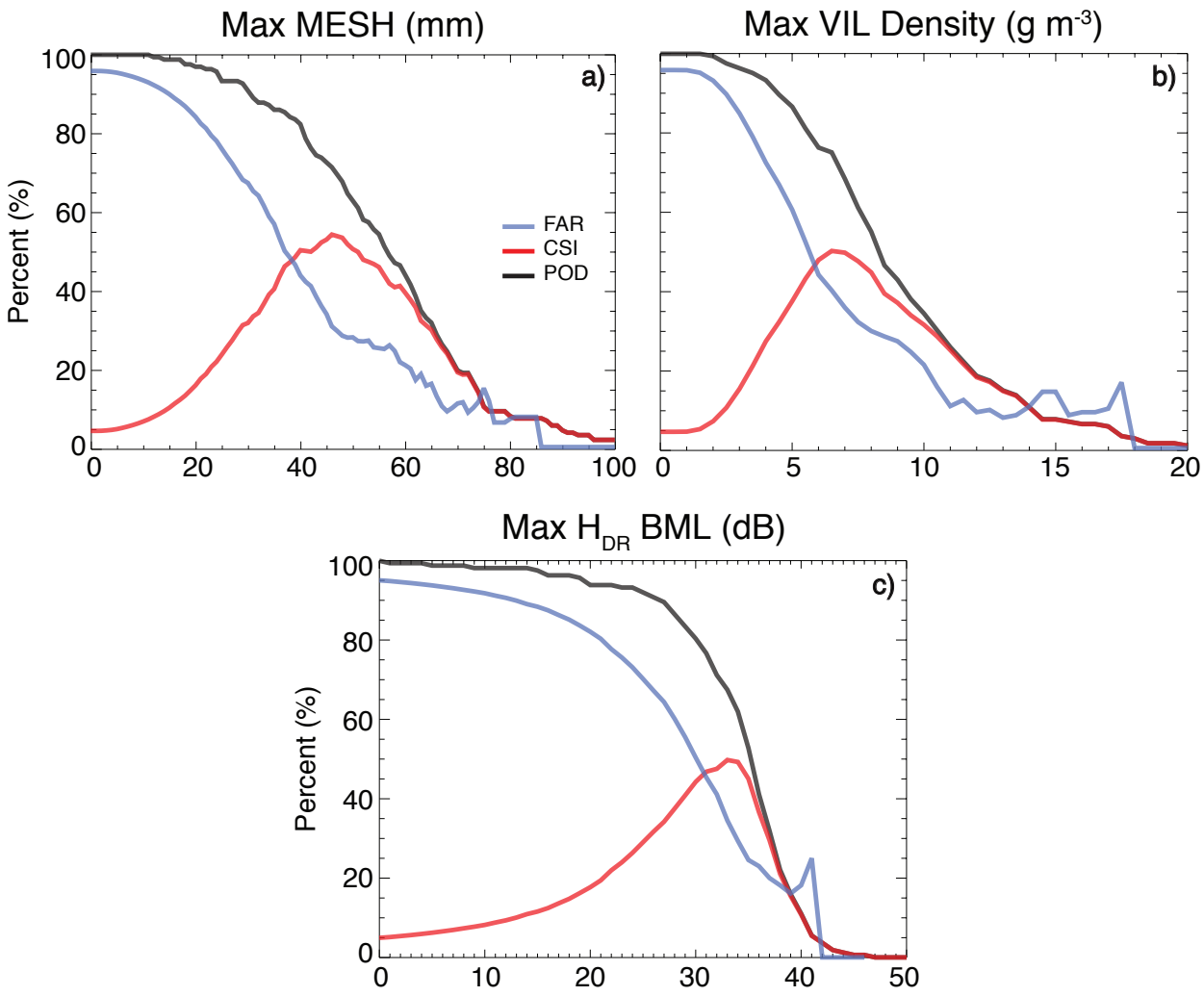

Figure 5.7: As in Fig. 5.5, but for significant hail-producing storms.

with H_{DR} BML (right panel, Figure 5.6) provides an approximate 5 percentage point increase in CSI relative to H_{DR} BML alone. Thus, using both SP and DP metrics appears to provide both statistical and practical increases in skill. It is also worth noting that the peak CSI tends to shift toward higher POD in the multiple variable approach. This performance evaluation, however, is sensitive to the sizes of each population. It is possible that the true skill is slightly lower than that depicted here, considering that we are minimizing the no-hail distribution to population dense regions only. Nonetheless, the relative performance is expected to be independent of these sample size limitations.

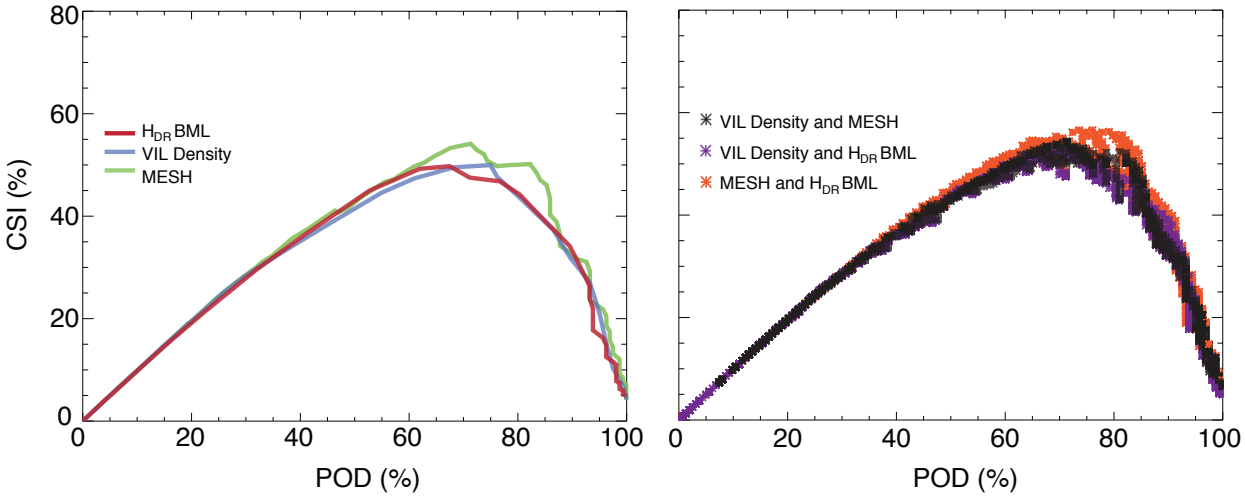

Figure 5.8: As in Fig. 5.6, but for significant hail-producing storms.

When evaluating the performance for only significant severe hail storms (those with hail exceeding 2 inches in diameter), CSI is lower and more variable across both the individual parameter and joint parameter space (Figs. 5.7 & 5.8). The increased variability in the performance metrics and decrease in overall skill is likely a product of decreased sample size for the hail storm population and a relative increase in the sample size for the null population, since the null population includes storms that produced smaller but potentially severe hail in this case. The location of the peak skill shifts to higher values of each metric in this analysis, indicating that each of these radar variables is correlated with hail size to some extent (as inferred from the analysis in Section 5.1). In contrast to the performance evaluation for all severe hail producing storms, the SP radar metrics slightly outperform H_{DR} BML for identification of significant severe hail storms. Substantial increases in skill and a shift of high skill to higher POD over the top performing single-variable method (MESH in this case) are still found when using a multiple variable method that combines SP and DP metrics for significant severe hail storms, especially for the combination of MESH and H_{DR} BML.

5.3 Objective hail size identification

As noted in Chapter 1, hail report sizes are biased towards reference objects, creating a discretized distribution of hail sizes. To address the discretization, uniform random noise was introduced to the report observations, producing small deviations in size. Using methods discussed in Allen et al. (2017), the relationship between the added noise, old, and new hail sizes are as follows:

$$y_{new} = y_{old} + y_{noise} \tag{5.1}$$

$$y_{noise} = \pm(0.247 y_{old} + 0.0279) \tag{5.2}$$

where y_{noise} is designed to avoid a large bias in small sizes and is capped at $\leq \pm$ 0.5 inches. This creates a more continuous, but similar shape, distribution of hail sizes that is more representative of what occurs in nature rather than as a result of reporting biases (see scatterplot distributions in Figure 5.10). The resulting report sizes after introducing uniform random noise is referred to as the smoothed hail sizes.

When comparing parameters to raw reported hail size, there is a lack of linearity evident (not shown). However, broad linear relationships were found between many parameters and smoothed hail size, mainly seen in higher parts of their distributions (Figure 5.9). Though the smoothed hail size distributions produced more linear relationships, the improvements are minimal. The lack of linearity for either report type between 25th and 50th percentiles for all three variables indicates the wide range of parameter values possible for a given reported hail size. This could be due to storms producing a range of hail sizes at a given time, rather than only the size that is recorded. However, the predicted MESH values based on the 75th percentile hail size equation in Witt et al. (1998a) had some considerable biases. Namely, MESH from Witt et al. (1998a) resulted in an underestimate of smaller hail sizes and an overestimate of larger hail sizes compared to the 75th percentile of the distribution of 5,954 hail reports analyzed here (Figure 5.10). It is important to note that this power law

relationship between SHI and MESH from Witt et al. (1998a) was empirically derived using a much smaller dataset than that in this study (147 hail reports). Thus, a revised power law relationship seems appropriate given the fact that MESH is routinely used in operational and climatological studies. The raw and smoothed report sizes were both tested to first assess the effect of adding randomized noise when fitting the distribution to percentile values. To that end, four revised MESH equations are presented in Figure 5.10 based on power law fits to the raw and smoothed report dataset based on the 75th percentile of hail sizes for each and the 95th percentile of hail sizes for each. When utilizing the raw report dataset, percentile bin values are constrained by the discretized report sizes, affecting the resulting power law functions. However, bins calculated using the smoothed report dataset follow a curve more similar to the distribution shape, resulting in a better power law fit. For this reason, we use the functions obtained using the smoothed report dataset going forward.

The vast majority of hail reports coincided with SHI values in a pattern consistent with the power law relationship, such that increased size was associated with increased SHI; however, 57 reports of 4.5-in hail fell outside the envelope in which nearly all other observations were located. Including these 57 reports skewed the 95th percentile calculations and the smoothed report dataset away from the distribution formed from the remaining 5,897 reports. Therefore, the 57 reports were flagged as erroneous and were neglected when introducing random noise (shown in red in Figure 5.10). The new MESH-SHI relationships for the 75th and 95th percentiles are as follows:

$$MESH_{75} = 16.566(SHI)^{0.181} \tag{5.3}$$

$$MESH_{95} = 17.270(SHI)^{0.272} \tag{5.4}$$

While other aspects of the Witt et al. (1998a) hail detection algorithm framework have been revised in previous work, including probability of severe hail (POSH; e.g. see Mahale

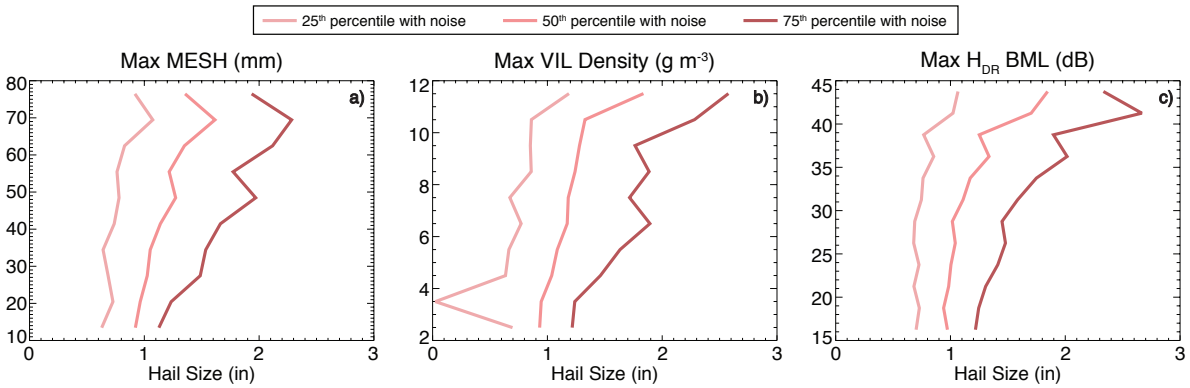

Figure 5.9: The 25th, 50th, and 75th percentiles of smoothed hail sizes via random noise for storm-maximum a) MESH, b) VIL density, and c) H_{DR} below the environmental melting level.

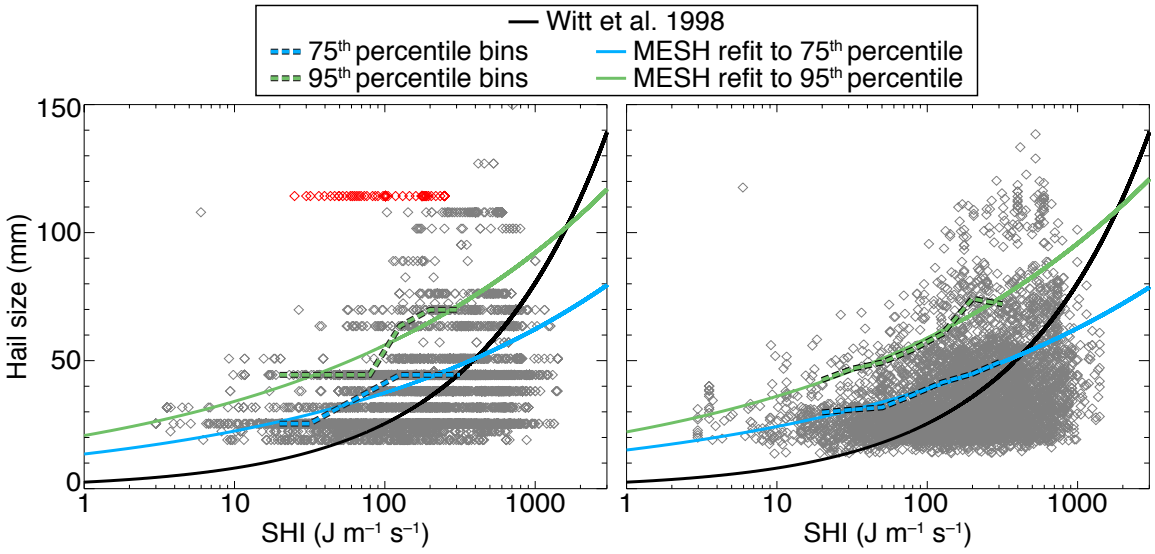

Figure 5.10: Scatterplot of SHI for all raw reported hail sizes (left) and smoothed hail sizes via random noise (right). Superimposed are the original MESH fit to the 75th percentile of hail size from Witt et al. (1998a) (gray line) and two revised MESH function fits to the 75th and 95th percentiles of hail size based on the data analyzed in this study (light and dark blue lines, respectively). The log-scale bin values used to constrain the new MESH fits are shown by the dashed lines. Reports colored red were excluded from the fit calculation for reasons discussed in Section 5.3.

et al. 2014), this is the first time (to the authors' knowledge) that the MESH-SHI relationship has been revisited.

To evaluate the practical utility of $MESH_{75}$ and $MESH_{95}$ in comparison to the original Witt et al. (1998a) relationship ($MESH_{WittEtAl}$ hereafter), we evaluated time-accumulated maps of MESH extremes (commonly referred to as "hail swaths") and coincident hail reports for several events analyzed in this study and provide two representative examples in Figure 5.11. The two events shown are 4 April 2017 over the southeast U.S. and 16 May 2017 over the central U.S. Great Plains. Hail swaths are limited to MESH values that at least meet severe hail criteria ($\geq$ 25.4 mm), to best evaluate severe hail false alarm and identification. These comparisons demonstrate that the new $MESH_{75}$ and $MESH_{95}$ relationships provide improved spatial coverage of areas where severe hail reports occurred and greater agreement between the largest predicted and observed hailstones compared to the $MESH_{WittEtAl}$ relationship, with $MESH_{95}$ providing the best overall estimate of maximum hail size. Though over-forecasting of both spatial coverage and stone size may be possible when using $MESH_{95}$, previous studies have documented the unrepresentative spatial coverage of hail reports, especially in areas of low population density (e.g. Allen and Tippett 2015), and have shown that hail reports tend to underestimate the true maximum hail size that reaches the surface (Bardsley 1990; Blair and Leighton 2012; Blair et al. 2017; Allen et al. 2017, and references therein). Furthermore, difficulty remains in providing adequate validations of radar-estimated hail fall given the broader limitations of hail reporting outlined in Chapter 1. It is worth noting here that we repeated the performance evaluations from Section 5.2 using the new $MESH_{75}$ and $MESH_{95}$ relationships and found no significant differences apart from changes in the MESH thresholds where peak skill was achieved. This result is not surprising given the fact that the only difference between MESH calculations is the coefficients of the power law fit.

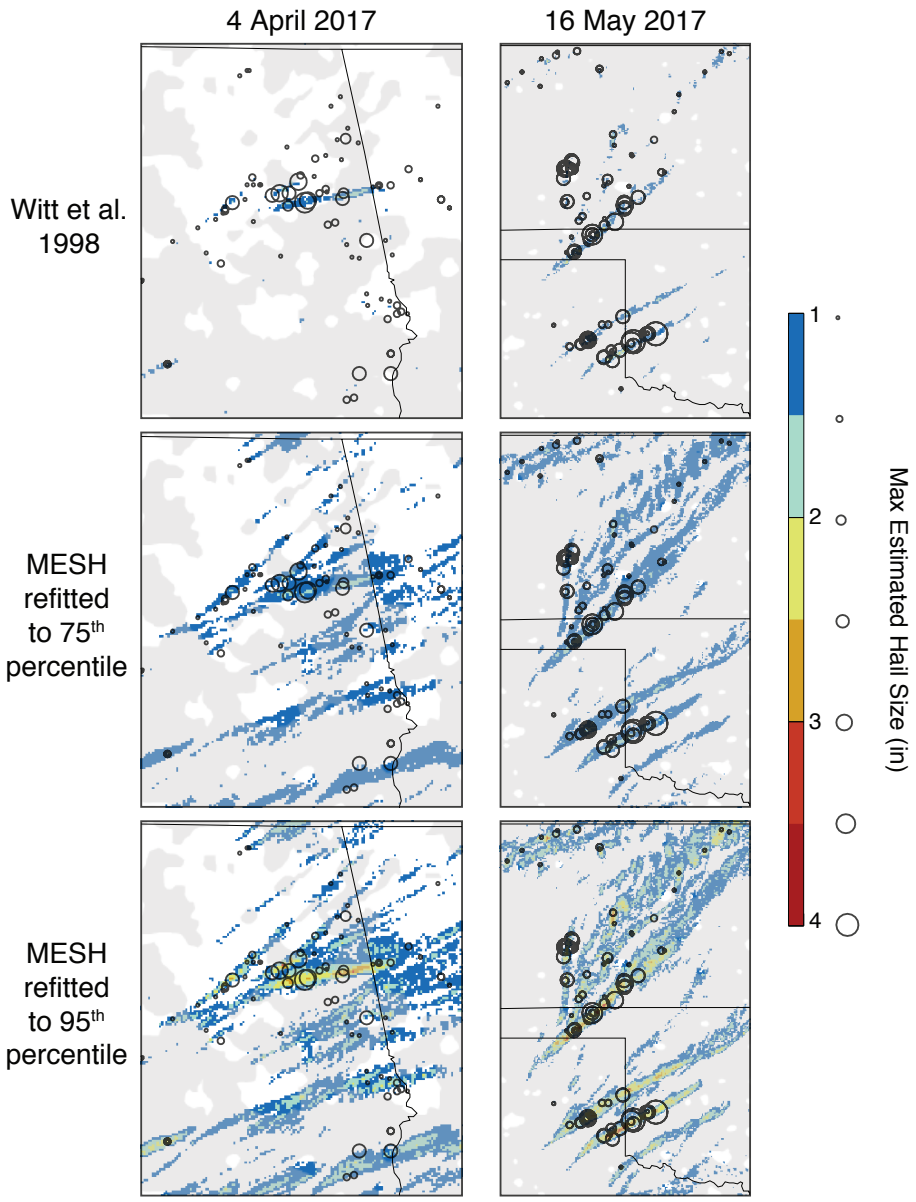

Figure 5.11: Time-accumulated maximum MESH values ("swaths"), limited to $\geq$ 25.4, over portions of the domains for (left) 4 April 2017 in the southeast U.S., and (right) 16 May 2017 in the central Great Plains. From top to bottom, maps are provided for the original Witt et al. (1998a) MESH equation, the revised fit to the 75th, and the 95th percentiles of observed hail sizes analyzed in this study. Black open circles indicate hail reports, which increase in diameter with increasing hail size (illustrated above the color bar at the bottom of the figure). Gray mask indicates regions that are less than 25 people per square mile.

Chapter 6

AACP-Severe Weather Linkage Results

6.1 Severe Weather Relationship

From a qualitative perspective, Figure 6.1 shows there are often severe weather reports clustered along AACP tracks, especially during the 11 May 2014 and 16 May 2017 events. The AACP tracks are normally longer than non-AACP storms, typical of supercells. The 5 April 2017 case is the most anomalous of the four, in that there are many long-lived AACP storms in the Southeast U.S. that were associated with very few or no reports. For example, 10 out of 25 AACP storms lasting $\geq$ 90 minutes had $<$ 5 reports. This area was within a NOAA SPC moderate risk for severe weather. Georgia and South Carolina were within a high risk, which is used to define regions with thermodynamic and wind shear conditions extremely favorable for severe weather. While many long-lived storms with intense updrafts occurred across this region, confirming the SPC risk assessment, it was peculiar that there were so few reports from these storms in such a favorable environment. There were several hail and wind reports from non-AACP storms across Alabama, Tennessee, and Kentucky on this day, exemplifying that AACPs cannot be used to detect all severe weather.

Severe weather relationships for the AACP and non-AACP storm populations are summarized in Table 6.1. 59% of AACP storms generated some form of severe weather. 78% of the severe weather reports occurred while an AACP was actively being produced, with the remaining 22% being split nearly evenly between time periods before or after AACP production (not shown). Hail (Tornado) was most (least) likely to be produced by an AACP storm. The overall low frequency of tornadoes explains this relationship, because, in general, only a small population of storms with evidence of strong rotating updrafts from

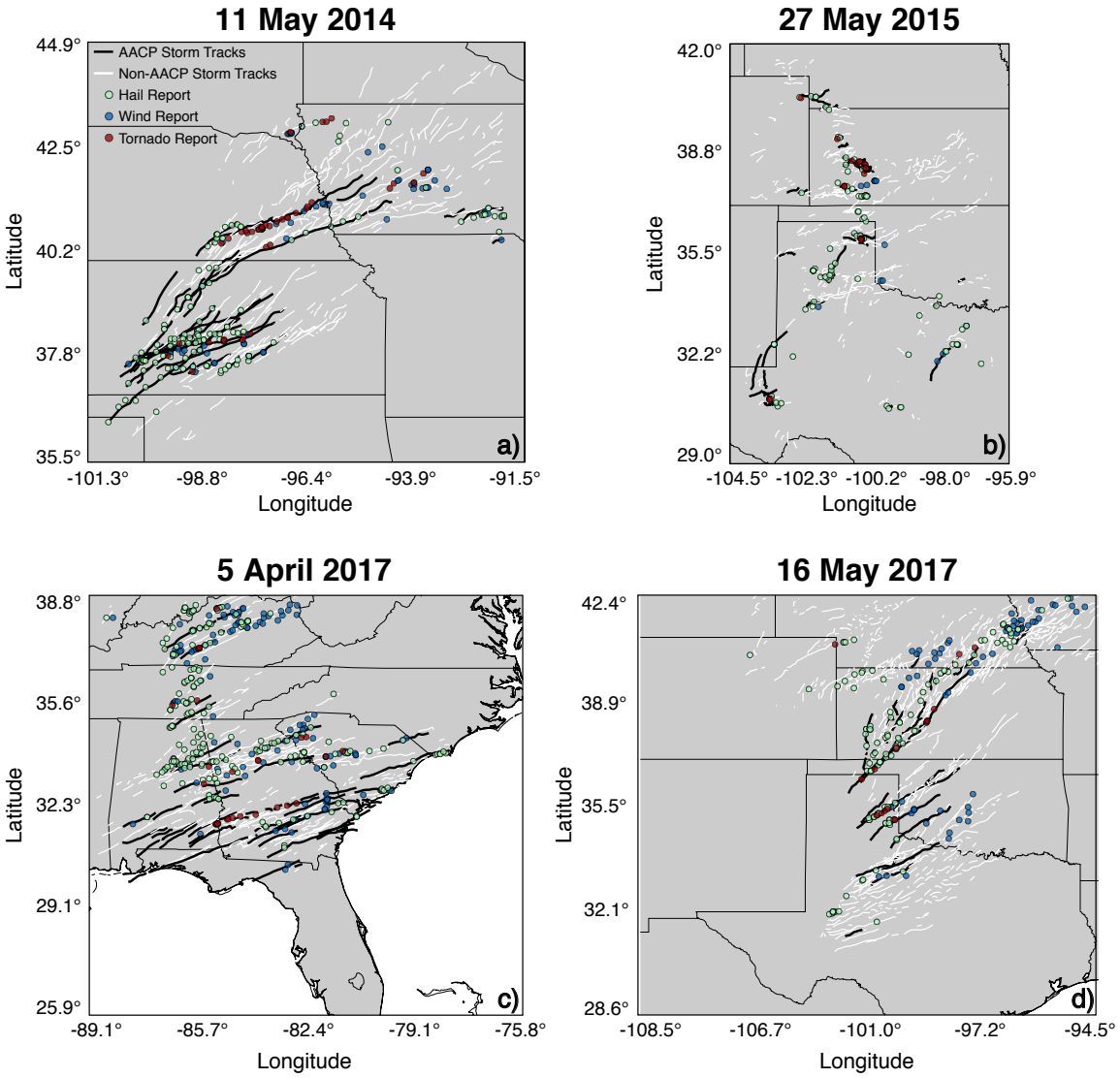

Figure 6.1: GridRad AACP (black) and non-AACP (white) storm tracks for 11-12 May 2014 (a), 27-28 May 2015 (b), 5-6 April 2017 (c), and 16-17 May 2017 (d), overlaid with severe weather reports (colored circles, see legend in upper-left panel).

Table 6.1: Severe weather characteristics of AACP storms, differentiated by severe weather type. Note that the fraction of AACP storms that produced hail, wind, or tornado (Column 2) should not average out to the fraction for any combination of severe.

	Fraction of AACP storms that produced the severe report	Fraction (No.) of severe reports linked to AACP storms	Fraction (No.) of significant (SIG) severe reports linked to AACP storms
Any Severe Report	59%	57% (4503)	73% (807)
Hail	48%	63% (3201)	88% (522)
Tornado	16%	68% (272)	86% (28)
Wind	22%	36% (1030)	41% (257)

Doppler radar produce a tornado. The relatively small fraction of AACP storms compared to the total (405 out of 4583 storms) generated 57% of all severe weather reports. If normalized by the number of storms in each category, AACP storms were responsible for 14 times the number of severe weather reports compared to non-AACP storms (6.33 vs 0.46 reports per storm). Of course, the existence of a trackable 40 dBZ echo alone would not be a sufficient indicator of a severe storm, and inclusion of a large population of non-severe storms decreases the number of reports per storm in this normalization. When all non-severe storms are disregarded, AACP storms generated 2.6 times the number of reports per storm than non-AACP storms (10.8 vs 4.2 reports per severe storm).

Storms that produce significant severe weather are of greatest threat to public safety and property, and 73% of these reports are attributed to AACP storms in this study. Significant hail (88%) and tornadoes (86%) were most associated with AACP storms, consistent with previous knowledge that supercells often generate AACPs and these hazards (Duda and Gallus 2010, and references therein). The relationship between AACPs and significant severe wind is much lower (41%), which may be attributed to the fact that 1) MCSs are

often responsible for the majority of significant wind reports, 2) AACPs were found to be shorter lived in MCSs than in discrete cells, and 3) the short AACP lifetime and short-lived nature of GridRad cells within MCSs reduced the quality of AACP-severe weather relationships. Trapp et al. (2006) found that severe wind reports can be biased and/or misleading which could also contribute to the reduced AACP - significant severe wind relationship.

Given that 88% of significant hail (2+ inch) reports were attributed to AACP storms, and severe weather most often occurs while an AACP is actively generated, knowledge that an AACP is being generated could increase NWS forecaster confidence that a given storm will produce significant hail. During the 13 severe weather events studied for AACP evaluation, there were 2211 severe thunderstorm or tornado warnings issued by the NWS within the domains observed by GOES super rapid scanning, 1846 of which feature a "tag" indicating the maximum expected size of hail. In total, 2338 storms were tagged to produce significant hail (127 warnings had more than one storm within the warning polygon). The number of significant hail events included in this analysis increased from 522 (see 2.8 and Table 6.1) to 765 because we removed the requirement that GOES imagery must be available for the entire GridRad storm cell lifetime.

We determine what fraction of NWS warnings with 2+inch hail tags were associated with significant severe hail, which defines POD. Conversely, we determine the fraction of warnings for which significant hail was forecast but no significant hail was reported, which defines FAR. The POD and FAR are derived for a combined set of warnings: 1) those with a 2+ inch hail tag and 2) those warnings with a 2+ inch hail tag or with an AACP. If an AACP was actively generated during a warning, we assume for this analysis that 2+ inch hail was predicted. POD and FAR statistics and the sample sizes (see equations 3.1-3.3) used to compute these statistics are provided in Table 6.2. 13.9% of significant hail events were captured by NWS warnings with 2+ inch hail tags, yet 75.3% of these tagged warnings were false alarms, i.e., 2+ inch hail never occurred despite being tagged. Warnings during

Table 6.2: POD, FAR, and CSI for storms with NWS warnings with a 2+ inch hail tag and any warning coincident with AACP generation based on an analysis of 2211 NWS severe thunderstorm or tornado warnings across 13 severe weather events.

	NWS Warnings with 2+ inch tag	NWS Warnings with 2+ inch tag or AACP classification
POD $\left(\frac{A}{A+C}\right)$	0.139	0.985
FAR $\left(\frac{B}{A+B}\right)$	0.753	0.864
CSI $\left(\frac{A}{A+B+C}\right)$	0.098	0.136
No. warnings that were hits (A)	37	262
No. warnings that were false alarms (B)	113	1658
No. warnings with < 2 inch tag or without AACP coincident with 2+ inch hail (C)	229	4

AACP generation or with a 2+ inch tag, captured 98.5% of these reports. FAR increases by 11% when AACPs are included. The CSI increases by 0.04 because the substantial POD increase outweighs the relatively small FAR increase.

It is important to note that unreported large hail events or inaccurate hail size reports will bias the statistics, but these biases are inherent to all severe weather studies dependent on reports (see Chapter 1). It is likely that significant hail occurs more frequently than the reports suggest which would lower the FAR statistics, but there is no immediate way to quantify this effect. FAR could also be lowered if, soon after forecasters identified an AACP, they were to look at radar and environmental information to see if significant hail was a realistic possibility before inserting a 2+ inch tag. Future work should include comparisons of NWS warnings and AACP with radar-derived hail size estimates, possibly from those discussed in Chapter 5, to consider more spatially continuous hail swaths and to minimize biases from population density. Nevertheless, based on the sample of storms analyzed

here, the large increase in POD indicates that forecaster recognition of an AACP at the time of a warning could improve confidence that significant hail will occur and encourage them to increase their hail size tag.

6.2 AACPs and Supercells

Given that supercells deviate from the mean flow and often feature intense, rotating updrafts, it is not unreasonable to expect AACP storms to be supercells, as AACPs are formed from intense updrafts in regions with strong mean flow-deviant motion. We found that 48% of AACP storms were supercells based on methods described in 2.2. This is not surprising given that AACP storm environments feature high convective available potential energy (CAPE) and strong deep layer (0-6 km) storm relative wind shear that is also favorable for supercells (Figure 6.2). However, AACP storms and supercells also occur in very similar environments to severe non-AACP storms. As expected, non-severe, non-AACP storms occur in weaker CAPE and shear environments (black line, Figure 6.2). 145 of the 194 total supercells (75%) produced an AACP. The 25% of supercells without an AACP had a mean 10 dBZ echo-top height at the tropopause level, whereas 75% of supercells with an AACP featured a mean 1 km above the tropopause (not shown). Although there is some overlap in the two populations and some non-AACP supercells did penetrate the tropopause by 2 km, the non-AACP supercells typically did not penetrate deeply enough into the stratosphere to generate a discernible AACP.

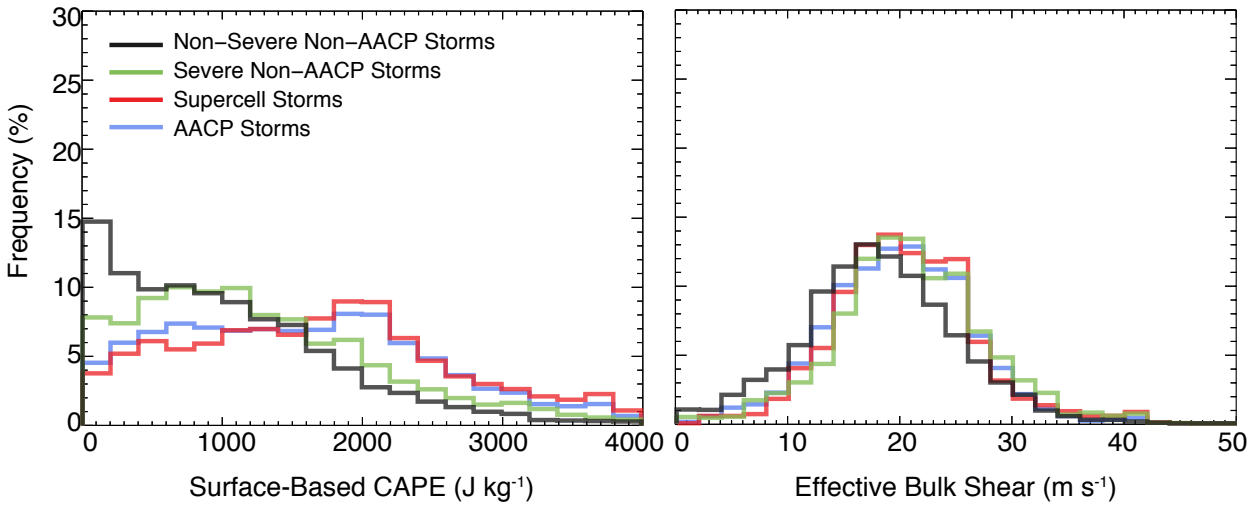

Figure 6.2: Distributions of surfaced-based CAPE (left) and 0-6 km effective bulk shear (right) for all four categories of storms.

Chapter 7

Conclusions

This study assessed two main objectives; the first of which was to analyze SP radar, DP radar, satellite, and lightning products to determine the best indicators for identifying hail events and hail-producing storms. Our conclusions from this analysis are as follows:

1. A population density threshold of 25 people per square mile along a storm's path appears to be a reliable criterion to establish a confident null population for hail storm analyses.

2. MESH, VIL density, and H_{DR} BML are the best discriminators for hail occurrence.

3. H_{DR} BML performs best for single parameter severe hail storm detection, while MESH performs best for significant severe hail storm detection.

4. A combination of SP and DP radar metrics (i.e., a multi-parameter approach) provides noticeably improved performance for hail storm detection compared to any single parameter approach.

5. Radar parameters that perform well at hail occurrence and hail storm identification scale roughly linearly with observed hail size, but no one parameter shows superior performance. However, revising the original MESH formulation from Witt et al. (1998a) to a relationship associated with the 95^{th} percentile of observed hail size improves the spatiotemporal comparisons between the hail reports and radar-based estimates.

The skill seen using MESH and VIL density to distinguish severe hail events is likely a result of two factors. Firstly, these parameters are computed by vertical integration throughout a storm. Larger values represent greater vertical depths, which correspond to

deeper/stronger convection that is more favorable for hail production. Secondly, Z_H, the variable that integrated, is most sensitive to particle size, such that large hydrometeors, especially hail, result in high Z_H values. H_{DR} BML is also skillful at hail identification due to the combined utility of Z_H and Z_{DR} that enables representation of physical characteristics within a storm. While Z_H provides information on particle size, Z_{DR} is sensitive to particle shape. Large hydrometeors that are also observed by the radar as relatively spherical (as hail is commonly seen due to tumbling motions as it falls) have different scattering characteristics than large, oblate rain drops. Therefore, combining Z_H and Z_{DR} results in skillful discrimination between hail and no-hail events. Conversely, those parameters that did not provide skillful discrimination either did not resolve the dynamical characteristics well enough to indicate major differences between the populations or did not adequately identify the physical characteristics of hydrometeors within a storm.

It is possible that the performance of discriminators evaluated here may change when evaluating single-radar data. The multi-radar composite GridRad data analyzed here are the result of distance- and time-weighted binning of single-radar observations (i.e., averaging), which results in slightly lower values of the radar variables than those seen in single-radar observations in most cases. Since all data is similarly impacted by this averaging, the relative performance of single-radar analyses is expected to be similar, with the threshold values at peak skill expected to increase in magnitude. Greater uncertainty is expected for metrics based on DP variables, since single-radar data are often impacted by systematic biases in Z_{DR} from poor radar calibration, which is a challenge for operational applications. Bias correction, such as the objective approach used when creating GridRad analyses, must be completed before any DP identification method is used.

We evaluated the performance of extreme values of hail metrics within radar-tracked storms rather than analyzing values coincident with the hail reports, which is often the approach taken in previous work. This choice was made in this study to avoid the spatiotemporal limitations of both the hail reports and the observations. While this approach

could result in unintended biases and negatively impacted performance, the results for several parameters are in agreement with past studies using the traditional space-time matching approach (including H_{DR} from Depue et al. (2007)).

The major findings in this study can be utilized to provide improved radar-based hail detection for future work. Potential applications include verification of model forecasts, hail climatologies using existing large-area radar data archives such as GridRad, the Multi-Radar Multi-Sensor system (MRMS) or its historical counterpart the Multi-Year Reanalysis of Remotely Sensed Storms (MYRORSS; Smith et al. 2016), and operational implementations (including real-time MRMS and potentially single-radar observations).

The second objective of this study was to quantify the relationship between AACPs and all severe weather. Our conclusions for this analysis are as follows:

1. On average, storms that produce an AACP feature a high frequency of severe weather, even when compared to other severe storms.

2. 75% of supercells featured an AACP and typically had deeper updrafts than the remaining 25%.

3. Generation of an AACP during an NWS warning showed a large increase in the likelihood that significant hail would occur, and knowledge of an AACP could improve upon warning-specified expected hail size.

In the warning process, forecasters are often considering multiple, complex datasets that provide near real-time inferences of in-storm dynamical and microphysical processes linked to severe weather. The NOAA NEXRAD radar network has gaps and deficiencies, so the existence of an AACP in under-observed regions can be used over the US and globally as a key decision aid for hail expectancy within a severe weather warning. AACPs can be identified in geostationary satellite imagery by anyone after a limited amount training, enabling a forecaster to quickly recognize these significant severe storms. This is especially true when 1-min or better imagery is available which allows one to best see the sometimes

subtle indicators of an AACP. The presence of an AACP could improve forecaster confidence to issue a warning in situations where radar signals are ambiguous or radar data is of reduced quality or unavailable. In addition, AACPs can be used to quickly identify the most intense storms, supercells in particular. Since AACP formation relies on strong updrafts and mean-flow-deviant motions, supercell storms often provide the necessary ingredients for them to occur. Monitoring AACP characteristics can help a forecaster identify updraft intensification or decay that may not be depicted well in a single volume of NEXRAD data available in routine NOAA NWS forecast operations. It would be advantageous to test these presumed benefits in an operational environment, such as the Hazardous Weather Test-bed Spring Forecasting Experiment conducted every summer in Norman, Oklahoma. Given that AACPs have also been documented throughout the world, and the majority is not observed by Doppler radar, the AACP-based severe storm identification capability could help save lives and protect property globally.

As mentioned in Section 2.1, limitations from AACP regional and seasonal sampling introduce some uncertainty as to how often AACPs occur in other regions/seasons across the US and the severity of these storms compared to the results described in section 6.1. AACP analyses in this thesis are limited to the Great Plains, Midwest and Southeast over only spring and summer, due to the availability of 1-minute satellite imagery. Supercell storms and extremely prominent overshooting updrafts (Cooney et al. 2018) are most common in these regions and in spring/summer, so we would expect AACPs to be most frequent in these regions. Storms with AACPs also certainly occur elsewhere and during other times of year according to Brunner et al. (2007). However, analyses of those regions/times could differ from those in this study. Additionally, we only include cases featuring mostly discrete convective modes, given the previously discussed difficulty of linking satellite features to radar storm tracks. The weaker observed relationship between severe wind and AACPs could likely be attributed to this exclusion of MCS cases, given that severe winds are most frequency associated with linear systems (Klimowski et al. 2003). Future work should be

devoted to analyzing significant severe weather occurrences in other regions, seasons, and convective modes to determine how often AACPs are associated with these events.

Lingering areas of uncertainty lie with interpretation of AACP IR temperature and microphysics, namely a lack of understanding of why some AACPs are warm but others are cold. This temperature and/or microphysical ambiguity will challenge a forecaster's ability to understand and identify AACPs. Convection-permitting models could be used to address this uncertainty, such as those described by Wang et al. (2016) and Homeyer et al. (2017) and references therein, provided they can generate realistic simulations of multiple AACPs with disparate temperature and microphysics. Although human analysts can identify AACPs with high accuracy, an automated AACP detection algorithm is required to determine the global distribution of AACP storms and extend AACP - severe weather analyses beyond a limited number of days. First-generation AACP detection methods developed within the GOES-R Algorithm Working Group (Bedka et al. 2011) and in Europe (Iršič Žibert and Žibert 2013) demonstrated that automated detection is possible, but there is opportunity to improve upon these methods using advanced pattern recognition approaches.

Bibliography

Allen, J., and M. Tippett, 2015: The characteristics of United States hail reports: 1955-2014. *Electron. J. Severe Storms Meteor.*, **10 (3)**, http://www.ejssm.org/ojs/index.php/ejssm/article/view/149/104.

Allen, J. T., M. K. Tippett, Y. Kaheil, A. H. Sobel, C. Lepore, S. Nong, and A. Muehlbauer, 2017: An extreme value model for U.S. hail size. *Mon. Wea. Rev.*, **145 (11)**, 4501–4519, doi:10.1175/MWR-D-17-0119.1.

Ambrosio, I. S., F. Martín, and F. Elizaga, 2007: Development and behaviour of a radar-based operational tool for hailstorms identification. *Atmos. Res.*, **83 (2)**, 306–314, doi: https://doi.org/10.1016/j.atmosres.2005.08.012.

Amburn, S. A., and P. L. Wolf, 1997: VIL density as a hail indicator. *Wea. Forecasting*, **12 (3)**, 473–478, doi:10.1175/1520-0434(1997)012⟨0473:VDAAHI⟩2.0.CO;2.

Apke, J. M., J. R. Mecikalski, and C. P. Jewett, 2016: Analysis of mesoscale atmospheric flows above mature deep convection using super rapid scan geostationary satellite data. *J. Appl. Meteor. Climatol.*, **55 (9)**, 1859–1887, doi:10.1175/JAMC-D-15-0253.1.

Aydin, K., T. A. Seliga, and V. Balaji, 1986: Remote sensing of hail with a dual linear polarization radar. *J. Climate Appl. Meteor.*, **25 (10)**, 1475–1484, doi:10.1175/1520-0450(1986)025⟨1475:RSOHWA⟩2.0.CO;2.

Bardsley, W. E., 1990: On the maximum observed hailstone size. *J. Appl. Meteor.*, **29 (11)**, 1185–1187, doi:10.1175/1520-0450(1990)029⟨1185:OTMOHS⟩2.0.CO;2.

Bedka, K., J. Brunner, R. Dworak, W. Feltz, J. Otkin, and T. Greenwald, 2010: Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients. *J. Appl. Meteor. Climatol.*, **49 (2)**, 181–202, doi: 10.1175/2009JAMC2286.1.

Bedka, K., J. Brunner, and W. Feltz, 2011: Objective overshooting top and enhanced-v signature detection for the goes-r advanced baseline imager: Algorithm theoretical basis document. Available online at: http://clouds.larc.nasa.gov/site/people/data/kbedka/GOES-R_ABI_ATBD_OvershootingTop_Enhanced-V_100perc.doc.

Bedka, K. M., and K. Khlopenkov, 2016: A probabilistic multispectral pattern recognition method for detection of overshooting cloud tops using passive satellite imager observations. *J. Appl. Meteor. Climatol.*, **55 (9)**, 1983–2005, doi:10.1175/JAMC-D-15-0249.1.

Bedka, K. M., and J. R. Mecikalski, 2005: Application of satellite-derived atmospheric motion vectors for estimating mesoscale flows. *J. Appl. Meteor.*, **44 (11)**, 1761–1772, doi:10.1175/JAM2264.1.

Bedka, K. M., C. S. Velden, R. A. Petersen, W. F. Feltz, and J. R. Mecikalski, 2009: Comparisons of satellite-derived atmospheric motion vectors, rawinsondes, and NOAA wind profiler observations. *J. Appl. Meteor. Climatol.*, **48 (8)**, 1542–1561, doi:10.1175/2009JAMC1867.1.

Bedka, K. M., C. Wang, R. Rogers, L. D. Carey, W. Feltz, and J. Kanak, 2015: Examining deep convective cloud evolution using total lightning, wsr-88d, and goes-14 super rapid scan datasets. *Wea. Forecasting*, **30 (3)**, 571–590, doi:10.1175/WAF-D-14-00062.1.

Blair, S. F., and J. W. Leighton, 2012: Creating high-resolution hail datasets using social media and post-storm ground surveys. *Electron. J. Severe Storms Meteor.*, **13**, 32–45.

Blair, S. F., and Coauthors, 2017: High-resolution hail observations: Implications for NWS warning operations. *Wea. Forecasting*, **32 (3)**, 1101–1119, doi:10.1175/WAF-D-16-0203.1.

Bringi, V. N., and V. Chandrasekar, 2001: *Polarimetric Doppler Weather Radar*. Cambridge University Press, 636 pp.

Bringi, V. N., J. Vivekanandan, and J. D. Tuttle, 1986: Multiparameter radar measurements in Colorado convective storms. part II: Hail detection studies. *J. Atmos. Sci.*, **43 (22)**, 2564–2577, doi:10.1175/1520-0469(1986)043⟨2564:MRMICC⟩2.0.CO;2.

Brotzge, J., S. Erickson, and H. Brooks, 2011: A 5-yr climatology of tornado false alarms. *Wea. Forecasting*, **26**, 534–544, doi:https://doi.org/10.1175/WAF-D-10-05004.1.

Browning, K. A., and J. Donaldson, 1963: Airflow and structure of a tornadic storm. *J. Atmos. Sci.*, **20 (6)**, 533–545.

Brunner, J. C., S. A. Ackerman, A. S. Bachmeier, and R. M. Rabin, 2007: A quantitative analysis of the enhanced-V feature in relation to severe weather. *Wea. Forecasting*, **22 (4)**, 853–872, doi:10.1175/WAF1022.1.

Bunkers, M. J., M. R. Hjelmfelt, and P. L. Smith, 2006: An observational examination of long-lived supercells. part i: Characteristics, evolution, and demise. *Wea. Forecasting*, **21 (5)**, 673–688, doi:10.1175/WAF949.1.

Capozzi, V., E. Picciotti, V. Mazzarella, F. S. Marzano, and G. Budillon, 2018: Fuzzy-logic detection and probability of hail exploiting short-range X-band weather radar. *Atmos. Res.*, **201**, 17–33, doi:https://doi.org/10.1016/j.atmosres.2017.10.006.

Cheng, L., M. English, and R. Wong, 1985: Hailstone size distributions and their relationship to storm thermodynamics. *J. Climate Appl. Meteor.*, **24 (10)**, 1059–1067, doi:10.1175/1520-0450(1985)024⟨1059:HSDATR⟩2.0.CO;2.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and A. K. Heidinger, 2013: Evolution of severe and nonsevere convection inferred from GOES-derived cloud properties. *J. Appl. Meteor. Climatol.*, **52 (9)**, 2009–2023, doi:10.1175/JAMC-D-12-0330.1.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29 (3)**, 639–653, doi:10.1175/WAF-D-13-00113.1.

Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous United States. *Wea. Forecasting*, **27 (5)**, 1235–1248, doi:10.1175/WAF-D-11-00151.1.

Cook, B., 1958: Hail determinatin by radar analysis. *Mon. Wea. Rev.*, **86 (11)**, 435–438, doi:10.1175/1520-0493(1958)086⟨0435:HDBRA⟩2.0.CO;2.

Cooney, J. W., K. P. Bowman, C. R. Homeyer, and T. M. Fenske, 2018: Ten year analysis of tropopause-overshooting convection using GridRad data. *J. Geophys. Res. Atmos.*, **123**, 329—-343, doi:https://doi.org/10.1002/2017JD027718.

Crum, T. D., and R. L. Alberty, 1993: The wsr-88d and the wsr-88d operational support facility. *Bull. Amer. Meteor. Soc.*, **74 (9)**, 1669–1688, doi:10.1175/1520-0477(1993)074⟨1669:TWATWO⟩2.0.CO;2.

Deierling, W., and W. A. Petersen, 2008: Total lightning activity as an indicator of updraft characteristics. *J. Geophys. Res.*, **113 (D16)**, doi:10.1029/2007JD009598.

Depue, T. K., P. C. Kennedy, and S. A. Rutledge, 2007: Performance of the hail differential reflectivity (H_{DR}) polarimetric radar hail indicator. *J. Appl. Meteor. Climatol.*, **46 (8)**, 1290–1301, doi:10.1175/JAM2529.1.

Donavon, R. A., and K. A. Jungbluth, 2007: Evaluation of a technique for radar identification of large hail across the upper midwest and central plains of the united states. *Wea. Forecasting*, **22 (2)**, 244–254, doi:10.1175/WAF1008.1.

Doswell, C. A. I., H. E. Brooks, and M. P. Kay, 2005: Climatological estimates of daily local nontornadic severe thunderstorm probability for the United States. *Wea. Forecasting*, **20 (4)**, 577–595, doi:10.1175/WAF866.1.

Doswell, C. A. I., and D. W. Burgess, 1993: Tornadoes and tornadic storms: A review of conceptual models. *Amer. Geophys. Union*, 161–172, doi:0.1029/GM079p0161.

Doviak, R. J., and D. S. Zrnić, 1993: *Doppler Radar and Weather Observations*. 2nd ed., Dover Publications, 562 pp.

Duda, J. D., and W. A. J. Gallus, 2010: Spring and summer midwestern severe weather reports in supercells compared to other morphologies. *Wea. Forecasting*, **25 (1)**, 190–206, doi:10.1175/2009WAF2222338.1.

Dworak, R., K. Bedka, J. Brunner, and W. Feltz, 2012: Comparison between GOES-12 overshooting-top detections, WSR-88D radar reflectivity, and severe storm reports. *Wea. Forecasting*, **27 (3)**, 684–699, doi:10.1175/WAF-D-11-00070.1.

Edwards, R., and R. L. Thompson, 1998: Nationwide comparisons of hail size with wsr-88d vertically integrated liquid water and derived thermodynamic sounding data. *Weather and Forecasting*, **13 (2)**, 277–285, doi:10.1175/1520-0434(1998)013⟨0277:NCOHSW⟩2.0.CO;2.

Elmore, K. L., 2011: The NSSL hydrometeor classification algorithm in winter surface precipitation: Evaluation and future development. *Wea. Forecasting*, **26 (5)**, 756–765, doi:10.1175/WAF-D-10-05011.1.

Farnell, C., T. Rigo, and N. Pineda, 2018: Exploring radar and lightning variables associated with the lightning jump. can we predict the size of the hail? *Atmos. Res.*, **202**, 175–186, doi:https://doi.org/10.1016/j.atmosres.2017.11.019.

Forbes, G. S., 1981: On the reliability of hook echoes as tornado indicators. *Mon. Wea. Rev.*, **109 (7)**, 1457—-1466.

Fraile, R., A. Castro, and J. Sánchez, 1992: Analysis of hailstone size distributions from a hailpad network. *Atmos. Res.*, **28 (3)**, 311–326, doi:https://doi.org/10.1016/0169-8095(92)90015-3.

Fujita, T. T., 1958: Mesoanalysis of the illinois tornadoes of 9 April 1953. *J. Meteor.*, **15**, 288–296.

Fujita, T. T., 1974: Overshooting thunderheads observed from ATS and learjet. Satellite and Mesometeorology Research Project Rep. 117, Texas Tech University, Lubbock, TX, 29 pp.

Fujita, T. T., 1982: Principle of stereographic height computations and their application to stratospheric cirrus over severe thunderstorms. *J. Meteorol. Soc. Japan*, **60**, 355—-368.

Fujita, T. T., D. L. Bradbury, and C. F. V. Thullenar, 1970: Palm sunday tornadoes of April 11, 1965. *Mon. Wea. Rev.*, **98 (1)**, 29–69, doi:10.1175/1520-0493(1970)098⟨0029:PSTOA⟩2.3.CO;2.

Goodman, S. J., and Coauthors, 2013: The GOES-R geostationary lightning mapper (GLM). *Atmos. Res.*, **125–126**, 34–49, doi:10.1016/j.atmosres.2013.01.006.

Gravelle, C. M., J. R. Mecikalski, W. E. Line, K. M. Bedka, R. A. Petersen, J. M. Sieglaff, G. T. Stano, and S. J. Goodman, 2016: Demonstration of a GOES-R Satellite convective toolkit to "bridge the gap" between severe weather watches and warnings: An example from the 20 May 2013 Moore, Oklahoma, tornado outbreak. *Bull. Amer. Meteor. Soc.*, **97 (1)**, 69–84, doi:10.1175/BAMS-D-14-00054.1.

Heinselman, P. L., and A. V. Ryzhkov, 2006: Validation of polarimetric hail detection. *Wea. Forecasting*, **21 (5)**, 839–850, doi:10.1175/WAF956.1.

Helmus, J., and S. Collis, 2016: The Python ARM Radar Toolkit (Py-ART), a Library for Working with Weather Radar Data in the Python Programming Language. *J. Open Research Software*, **4 (1)**, p.e25, doi:http://doi.org/10.5334/jors.119.

Herzegh, P. H., and A. R. Jameson, 1992: Observing precipitation through dual-polarization radar measurements. *Bull. Amer. Meteor. Soc.*, **73 (9)**, 1365–1376, doi:10.1175/1520-0477(1992)073⟨1365:OPTDPR⟩2.0.CO;2.

Heymsfield, A. J., I. M. Giammanco, and R. Wright, 2014: Terminal velocities and kinetic energies of natural hailstones. *Geophys. Res. Lett.*, **41 (23)**, 8666–8672, doi:10.1002/2014GL062324.

Heymsfield, G., R. Blackmer, and S. Schotz, 1983: Upper-level structure of oklahoma tornadic storms on 2 May 1979. I: radar and satellite observations. *J. Atmos. Sci.*, **40**, 740––1755, doi:https://doi.org/10.1175/1520-0469(1983)040⟨1740:ULSOOT⟩2.0.CO;2.

Holleman, I., H. Wessels, J. Onvlee, and S. Barlag, 2000: Development of a hail-detection-product. *Phys. Chem. Earth. 25B*, **25 (10)**, 1293–1297, doi:https://doi.org/10.1016/S1464-1909(00)00197-0.

Homeyer, C. R., 2014: Formation of the enhanced-V infrared cloud-top feature from high-resolution three-dimensional radar observations. *J. Atmos. Sci.*, **71 (1)**, 332–348, doi:10.1175/JAS-D-13-079.1.

Homeyer, C. R., and K. P. Bowman, 2017: Algorithm Description Document for Version 3.1 of the Three-Dimensional Gridded NEXRAD WSR-88D Radar (GridRad) Dataset. Available online at: http://gridrad.org/pdf/GridRad-v3.1-Algorithm-Description.pdf.

Homeyer, C. R., and M. R. Kumjian, 2015: Microphysical characteristics of overshooting convection from polarimetric radar observations. *J. Atmos. Sci.*, **72 (2)**, 870–891, doi:10.1175/JAS-D-13-0388.1.

Homeyer, C. R., J. D. McAuliffe, and K. M. Bedka, 2017: On the development of above-anvil cirrus plumes in extratropical convection. *J. Atmos. Sci.*, **74 (5)**, 1617–1633, doi:10.1175/JAS-D-16-0269.1.

Hubbert, J., V. Bringi, L. Carey, and S. Bolen, 1998: Csu-chill polarimetric radar measurements from a severe hail storm in eastern colorado. *J. Appl. Meteor.*, **37**, 749–775.

Iowa State University, 2018: https://mesonet.agron.iastate.edu/request/gis/watchwarn.phtml.

Iršič Žibert, M., and J. Žibert, 2013: Monitoring and automatic detection of the cold-ring patterns atop deep convective clouds using meteosat data. *Atmos. Res.*, **37**, 281—-292.

Kelly, D. L., J. T. Schaefer, and C. A. Doswell, 1985: Climatology of nontornadic severe thunderstorm events in the united states. *Mon. Wea. Rev.*, **113 (11)**, 1997–2014, doi:10.1175/1520-0493(1985)113⟨1997:CONSTE⟩2.0.CO;2.

Kennedy, P. C., S. A. Rutledge, W. A. Petersen, and V. N. Bringi, 2001: Polarimetric radar observations of hail formation. *J. Appl. Meteor.*, **40 (8)**, 1347–1366, doi:10.1175/1520-0450(2001)040⟨1347:PROOHF⟩2.0.CO;2.

Klimowski, B. A., M. J. Bunkers, M. R. Hjelmfelt, and J. N. Covert, 2003: Severe convective windstorms over the northern high plains of the united states. *Wea. Forecasting*, **18 (3)**, 502–519, doi:10.1175/1520-0434(2003)18⟨502:SCWOTN⟩2.0.CO;2.

Kumjian, M. R., 2013a: Principles and applications of dual-polarization weather radar. part i: Description of the polarimetric radar variables. *J. Oper. Meteor.*, **1**, 226–242, doi:10.15191/nwajom.2013.0119.

Kumjian, M. R., 2013b: Principles and applications of dual-polarization weather radar. part ii: Warm and cold season applications. *J. Oper. Meteor.*, **1**, 243–264, doi:10.15191/nwajom.2013.0120.

Kumjian, M. R., 2013c: Principles and applications of dual-polarization weather radar. part iii: Artifacts. *J. Oper. Meteor.*, **1**, 265–274, doi::10.15191/nwajom.2013.0121.

Kumjian, M. R., A. P. Khain, N. Benmoshe, E. Ilotoviz, A. V. Ryzhkov, and V. T. J. Phillips, 2014: The anatomy and physics of zdr columns: Investigating a polarimetric radar signature with a spectral bin microphysical model. *J. Appl. Meteor. Climatol.*, **53 (7)**, 1820–1843, doi:10.1175/JAMC-D-13-0354.1.

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47 (7)**, 1940–1961, doi:10.1175/2007JAMC1874.1.

Kunz, M., U. Blahak, J. Handwerker, M. Schmidberger, H. J. Punge, S. Mohr, E. Fluck, and K. M. Bedka, 2018: The severe hailstorm in southwest Germany on 28 July 2013: characteristics, impacts and meteorological conditions. *Quart. J. Roy. Meteor. Soc.*, **144 (710)**, 231–250, doi:10.1002/qj.3197.

Lemon, L. R., and C. A. I. Doswell, 1979: Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis. *Mon. Wea. Rev.*, **107 (9)**, 1184–1197.

Lindsey, D. T., D. W. Hillger, L. Grasso, J. A. Knaff, and J. F. Dostalek, 2006: Goes climatology and analysis of thunderstorms with enhanced 3.9-μm reflectivity. *Mon. Wea. Rev.*, **134 (9)**, 2342–2353, doi:10.1175/MWR3211.1.

Line, W. E., T. J. Schmit, D. T. Lindsey, and S. J. Goodman, 2016: Use of geostationary super rapid scan satellite imagery by the storm prediction center. *Wea. Forecasting*, **31 (2)**, 483–494, doi:10.1175/WAF-D-15-0135.1.

Liu, C., and S. Heckman, 2010: The application of total lightning detection and cell tracking for severe weather prediction. *TECO-2010-WMO Technical Conf. on Meteor. and Environmental Instruments and Methods of Observation*, Helsinki, Finland, 10 pp.

López, L., and J. Sanchez, 2009: Discriminant methods for radar detection of hail. *Atmos. Res.*, **93**, 358–368.

Mahale, V. N., G. Zhang, and M. Xue, 2014: Fuzzy logic classification of s-band polarimetric radar echoes to identify three-body scattering and improve data quality. *J. Appl. Meteor.*, **53 (8)**, 2017–2033, doi:10.1175/JAMC-D-13-0358.1.

Marzban, C., and A. Witt, 2001: A bayesian neural network for severe-hail size prediction. *Wea. Forecasting*, **16 (5)**, 600–610, doi:10.1175/1520-0434(2001)016⟨0600:ABNNFS⟩2.0.CO;2.

Mather, G. K., D. Treddenick, and R. Parsons, 1976: An observed relationship between the height of the 45 dBZ contours in storm profiles and surface hail reports. *J. Appl. Meteor.*, **15 (12)**, 1336–1340, doi:10.1175/1520-0450(1976)015⟨1336:AORBTH⟩2.0.CO;2.

McCann, D. W., 1983: The enhanced-V: A satellite observable severe storm signature. *Mon. Wea. Rev.*, **111 (4)**, 887–894, doi:10.1175/1520-0493(1983)111⟨0887:TEVASO⟩2.0.CO;2.

Menzel, W. P., and J. F. W. Purdom, 1994: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites. *Bull. Amer. Meteor. Soc.*, **75 (5)**, 757–782, doi:10.1175/1520-0477(1994)075⟨0757:IGITFO⟩2.0.CO;2.

NOAA/NCEI, 2013: Space weather data. NOAA/NCEI, accessed August 2016 to July 2018, https://www.ncdc.noaa.gov/data-access/satellite-data.

NOAA/NCEI, 2014: Storm events database, version 3. NOAA/NCEI, accessed August 2016 to July 2018, https://www.ncdc.noaa.gov/stormevents/.

NOAA/NCEP/ESRL, 2012: Rapid Refresh (RAP). NOAA/NCEP/ESRL, accessed August 2016 to July 2018, https://rapidrefresh.noaa.gov/.

NOAA/NESDIS/NCEI, 1991: NOAA Next Generation Radar (NEXRAD) Level II Base Data. NOAA/NESDIS/NCEI, accessed August 2016 to July 2018, doi:10.7289/V5W9574V.

NOAA/NWS/SPC, 1955: Severe Weather Database. NOAA/NWS/SPC, accessed August 2016 to July 2018, https://www.spc.noaa.gov/wcm/.

Ortega, K., 2018: Evaluating multi-radar, multi-sensor products for surface hailfall diagnosis. *Electron. J. Severe Storms Meteor.*, **13 (1)**, URL http://www.ejssm.org/ojs/index.php/ejssm/article/view/163/113.

Ortega, K. L., J. M. Krause, and A. V. Ryzhkov, 2016: Polarimetric radar characteristics of melting hail. part III: Validation of the algorithm for hail size discrimination. *J. Appl. Meteor. Climatol.*, **55 (4)**, 829–848, doi:10.1175/JAMC-D-15-0203.1.

Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. G. Kolodziej, and J. J. Gourley, 2009: The severe hazards analysis and verification experiment. *Bull. Amer. Meteor. Soc.*, **90 (10)**, 1519–1530, doi:10.1175/2009BAMS2815.1.

Ortega, K. L., T. M. Smith, G. J. Stumpf, J. Hocker, and L. Lopez, 2005: A comparison of multi-sensor hail diagnosis techniques. Preprints, *21st Conf. on Interactive Information Processing Systems*, San Diego, CA, Amer. Meteor. Soc., P1.11., [Available online at http://ams.confex.com/ams/pdfpapers/87640.pdf].

Park, H. S., A. V. Ryzhkov, D. S. Zrnić, and K.-E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, **24 (3)**, 730–748, doi:10.1175/2008WAF2222205.1.

Petrocchi, P. J., 1982: Automatic detection of hail by radar. AFGLTR-82-0277. Environmental Research Paper 796, Air Force Geophysics Laboratory, Hanscom AFB, MA, 33 pp. [Available from Air Force Geophysics Laboratory, Hanscom AFB, MA 01731].

Picca, J., and A. Ryzhkov, 2012: A dual-wavelength polarimetric analysis of the 16 may 2010 oklahoma city extreme hailstorm. *Mon. Wea. Rev.*, **140 (4)**, 1385–1403, doi:10.1175/MWR-D-11-00112.1.

Púčik, T., M. Valachová, and P. Zacharov, 2013: Upper tropospheric conditions in relation to the cloud top features of 15 august 2010 convective storms. *Atmos. Res.*, **123**, 249–267, doi:http://dx.doi.org/10.1016/j.atmosres.2012.10.006.

Rosenfeld, D., W. L. Woodley, A. Lerner, G. Kelman, and D. T. Lindsey, 2008: Satellite detection of severe convective storms by their retrieved vertical profiles of cloud particle effective radius and thermodynamic phase. *J. Geophys. Res.*, **113**, doi:10.1029/2007JD008600.

Ryzhkov, A. V., S. E. Giangrande, V. Melnikov, and T. Schuur, 2005a: Calibration issues of dual-polarization radar measurements. *J. Atmos. Oceanic. Technol.*, **22**, 1130–1155.

Ryzhkov, A. V., M. R. Kumjian, S. M. Ganson, and P. Zhang, 2013: Polarimetric radar characteristics of melting hail. part II: Practical implications. *J. Appl. Meteor. Climatol.*, **52 (12)**, 2871–2886, doi:10.1175/JAMC-D-13-074.1.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnić, 2005b: The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *J. Atmos. Oceanic. Technol.*, **86**, 809–824.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, and D. S. Zrnić, 2005c: Polarimetric tornado detection. *J. Appl. Meteor.*, **44 (5)**, 557–570, doi:10.1175/JAM2235.1.

Sandmæl, T. N., 2017: An evaluation of radar- and satellite-data based products to discriminate between tornadic and non-tornadic storms. School of Meteorology, University of Oklahoma, 98 pp., https://hdl.handle.net/11244/52775.

Schaefer, J., J. J. Levit, S. J. Weiss, and D. W. McCarthy, 2004: The frequency of large hail over the contiguous United States. *14th Conf. on Applied Climatology*, Seattle, WA, Amer. Meteor. Soc., 3.3, [Available online at https://ams.confex.com/ams/pdfpapers/69834.pdf].

Schmit, T. J., M. M. Gunshor, W. P. Menzel, J. J. Gurka, J. Li, and A. S. Bachmeier, 2005: Introducing the next-generation advanced baseline imager on GOES-R. *Bull. Amer. Meteor. Soc.*, **86 (8)**, 1079–1096, doi:10.1175/BAMS-86-8-1079.

Schmit, T. J., and Coauthors, 2014: GOES-14 super rapid scan operations to prepare for GOES-R. *J. Appl. Remote. Sens*, **7 (1)**, doi:10.1117/1.JRS.7.073462.

Schultz, C. J., L. D. Carey, E. V.Schultz, and R. J.Blakeslee, 2017: Kinematic and microphysical significance of lightning jump versus nonjump increases in total flash rate. *Wea. Forecasting*, **32 (1)**, 275–288, doi:10.1175/WAF-D-15-0175.1.

Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *J. Appl. Meteor. Climatol.*, **48 (12)**, 2543–2563, doi:10.1175/2009JAMC2237.1.

Setvák, M., K. M. Bedka, D. T. Lindsey, A. Sokol, Z. Charvát, J. Šťástka, and P. K. Wang, 2013: A-train observations of deep convective storm tops. *Atmos. Res.*, **123**, 229–248, doi:https://doi.org/10.1016/j.atmosres.2012.06.020.

Setvák, M., and Coauthors, 2010: Satellite-observed cold-ring-shaped features atop deep convective clouds. *Atmos. Res.*, **97 (1)**, 80–96, doi:https://doi.org/10.1016/j.atmosres.2010.03.009.

Skripniková, K., and D. Řezáčová, 2014: Radar-based hail detection. *Atmos. Res.*, **144**, 175–185, doi:https://doi.org/10.1016/j.atmosres.2013.06.002.

Smith, A. B., and R. W. Katz, 2013: US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. *Natural Hazards*, **67 (2)**, 387–410, doi:10.1007/s11069-013-0566-5.

Smith, A. B., and J. L. Matthews, 2015: Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Natural Hazards*, **77 (3)**, 1829–1851, doi:10.1007/s11069-015-1678-x.

Smith, T. M., and Coauthors, 2016: Multi-radar multi-sensor (mrms) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97 (9)**, 1617–1630, doi:10.1175/BAMS-D-14-00173.1.

Snyder, J. C., A. V. Ryzhkov, M. R. Kumjian, A. P. Khain, and J. Picca, 2015: A zdr column detection algorithm to examine convective storm updrafts. *Wea. Forecasting*, **30 (6)**, 1819–1844, doi:10.1175/WAF-D-15-0068.1.

Starzec, M., C. R. Homeyer, and G. L. Mullendore, 2017: Storm labeling in three dimensions (SL3D): A volumetric radar echo and dual-polarization updraft classification algorithm. *Mon. Wea. Rev.*, **145 (3)**, 1127–1145, doi:10.1175/MWR-D-16-0089.1.

Straka, J. M., D. S. Zrnić, and A. V. Ryzhkov, 2000: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. *J. Appl. Meteor.*, **39 (8)**, 1341–1372, doi:10.1175/1520-0450(2000)039⟨1341:BHCAQU⟩2.0.CO;2.

Strong, G., and E. Lozowski, 1977: An Alberta study to objectively measure hailfall intensity. *Atmosphere*, **15**, 33–53.

Trapp, R. J., G. Stumpf, and K. Manross, 2005: A reassessment of the percentage of tornadic mesocyclones. *Wea. Forecasting*, **20**, 680—-687, doi:https://doi.org/10.1175/WAF864.1.

Trapp, R. J., D. M. Wheatley, N. T. Atkins, R. W. Przybylinski, and R. Wolf, 2006: Buyer beware: Some words of caution on the use of severe wind reports in postevent assessment and research. *Wea. Forecasting*, **21 (3)**, 408–415, doi:10.1175/WAF925.1.

University of Wisconsin - Madison Space Science and Engineering Center, 2011: Space weather data. Accessed August 2016 to July 2018, https://www.ssec.wisc.edu/data/geo-list.

Verbout, S., H. Brooks, L. Leslie, and D. Schultz, 2006: Evolution of the u.s. tornado database. *Wea. Forecasting*, **21**, 86–93, doi:https://doi.org/10.1175/WAF910.1.

Vivekanandan, J., D. S. Zrnić, S. M. Ellis, R. Oye, A. V. Ryzhkov, and J. Straka, 1999: Cloud microphysics retrieval using S-Band dual-polarization radar measurements. *Bull. Amer. Meteorol. Soc.*, **80**, 381–388.

Waldvogel, A., B. Federer, and P. Grimm, 1979: Criteria for the detection of hail cells. *J. Appl. Meteor.*, **18 (12)**, 1521–1525, doi:10.1175/1520-0450(1979)018⟨1521:CFTDOH⟩2.0.CO;2.

Wang, P., J. Shi, J. Hou, and Y. Hu, 2018: The identification of hail storms in the early stage using time series analysis. *J. Geophys. Res.*, **123 (2)**, 929–947, doi:10.1002/2017JD027449.

Wang, P. K., 2003: Moisture plumes above thunderstorm anvils and their contributions to cross tropopause transport of water vapor in midlatitudes. *J. Geophys. Res.*, **108 (D6)**, doi:10.1029/2003JD002581.

Wang, P. K., K. Y. Cheng, M. Setvak, and C. K. Wang, 2016: The origin of the gullwing-shaped cirrus above an argentinian thunderstorm as seen in calipso images. *J. Geophys. Res. Atmos.*, **121**, 3729—-3738, doi:10.1002/2015JD024111.

Williams, E., and Coauthors, 1999: The behavior of total lightning activity in severe florida thunderstorms. *Atmos. Res.*, **51 (3)**, 245–265, doi:https://doi.org/10.1016/S0169-8095(99)00011-3.

Witt, A., D. Burgess, A. Seimon, J. T. Allen, J. Snyder, and H. Bluestein, 2018: Rapid-scan radar observations of an oklahoma tornadic hailstorm producing giant hail. *In Press, Weather and Forecasting*.

Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. W. Mitchell, and K. W. Thomas, 1998a: An enhanced hail detection algorithm for the WSR-88D. *Wea. Forecasting*, **13 (2)**, 286–303, doi:10.1175/1520-0434(1998)013⟨0286:AEHDAF⟩2.0.CO;2.

Witt, A., M. D. Eilts, G. J. Stumpf, E. D. W. Mitchell, J. T. Johnson, and K. W. Thomas, 1998b: Evaluating the performance of wsr-88d severe storm detection algorithms. *Wea. Forecasting*, **13 (2)**, 513–518, doi:10.1175/1520-0434(1998)013⟨0513:ETPOWS⟩2.0.CO;2.

Yost, C. R., and Coauthors, 2018: A prototype method for diagnosing high ice water content probability using satellite imager data. *Atmos. Meas. Tech.*, **11 (2)**, 1615—1637, doi:https://doi.org/10.5194/amt-11-1615-2018.

Zrnić, D. S., and A. V. Ryzhkov, 1999: Polarimetry for weather surveillance radars. *Wea. Forecasting*, **80**, 389–406, doi:0.1175/1520-0477(1999)080,0389:PFWSR.2.0.CO;2.